

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
22 September 2005 (22.09.2005)

PCT

(10) International Publication Number
WO 2005/086891 A2(51) International Patent Classification: **Not classified**(21) International Application Number:
PCT/US2005/007894

(22) International Filing Date: 7 March 2005 (07.03.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/550,810 5 March 2004 (05.03.2004) US
60/604,076 24 August 2004 (24.08.2004) US
60/650,401 4 February 2005 (04.02.2005) US(71) Applicants (for all designated States except US):
ROSETTA INPHARMATICS LLC [US/US]; 401 Terry Avenue North, Seattle, WA 98109 (US). **THE NETHERLANDS CANCER INSTITUTE** [NL/NL]; Plesmanlaan 121, NL-1066 CX Amsterdam (NL).

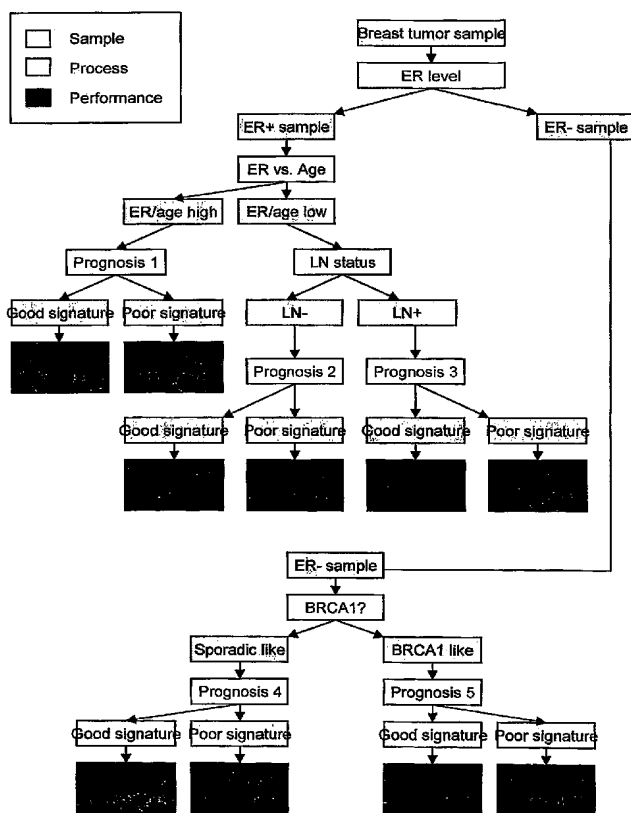
(72) Inventors; and

(75) Inventors/Applicants (for US only): **DAI, Hongyue**[CN/US]; 16814 118th Avenue NE, Bothell, WA 98011 (US). **VAN'T VEER, Laura, J.** [NL/NL]; Brouwersgracht 192-G, NL-1013 HC Amsterdam (NL). **LAMB, John** [GB/US]; 1216 N 172nd Street, Shoreline, WA 98133 (US). **STOUGHTON, Roland** [US/US]; Apt. D, 5919 Mildred Street, San Diego, CA 92110 (US). **FRIEND, Stephen, H.** [US/US]; 101 W. Mermaid Lane, Philadelphia, PA 19118 (US). **HE, Yudong** [US/US]; 11410 NE 124th Street #148, Kirkland, WA 98034 (US).(74) Agents: **ANTLER, Adriane, M.** et al.; Jones Day, 222 East 41st Street, New York, NY 10017-6702 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,

[Continued on next page]

(54) Title: CLASSIFICATION OF BREAST CANCER PATIENTS USING A COMBINATION OF CLINICAL CRITERIA AND INFORMATIVE GENESETS



(57) Abstract: The present invention provides prognostic methods for conditions such as cancer, for example, breast cancer, comprising classifying an individual by a plurality of phenotypic, genotypic or clinical characteristics of the condition into a plurality of patient subsets, and analyzing the pattern of expression of prognosis-informative genes identified for that subset in a sample from the individual. The present invention also provides methods for constructing such patient subsets and of identifying prognosis-informative genesets for such subsets. The invention further provides methods of assigning a therapeutic regimen to an individual, microarrays useful for performing prognosis, kits comprising these microarrays, and computer systems and programs for implementing the methods of the invention.

WO 2005/086891 A2



PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

(48) Date of publication of this corrected version:

26 January 2006

(15) Information about Correction:

see PCT Gazette No. 04/2006 of 26 January 2006, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

CLASSIFICATION OF BREAST CANCER PATIENTS USING A COMBINATION OF CLINICAL CRITERIA AND INFORMATIVE GENESETS

[0001] This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Patent Application No. 60/650,401, filed on February 4, 2005, U.S. Provisional Patent Application No. 60/604,076, filed on August 24, 2004, and U.S. Provisional Patent Application No. 60/550,810, filed on March 5, 2004, each of which is incorporated by reference herein in its entirety.

1. FIELD OF THE INVENTION

[0002] The present invention relates to the use of both phenotypic and genotypic aspects of a condition, such as a disease, in order to identify discrete subsets of patients for which specific sets of informative genes are then identified. The invention also relates to the classification of individuals, such as breast cancer patients, into a subset of the condition on the basis of clinical parameters and the status of markers, for example, of genes expression patterns, and the prognosis of those individuals on the basis of markers informative for prognosis within the subset of the condition. The invention also relates to methods of determining a course of treatment or therapy to an individual having, or suspected of having, a condition, such as breast cancer. The invention further relates to methods of structuring a clinical trial, particularly using five breast cancer-specific patient subsets and prognosis-informative genes for each, and of identifying patient populations for clinical trials or for other condition-related, for example, breast cancer-related, research. Finally, the invention relates to computer implementations of the above methods.

2. BACKGROUND OF THE INVENTION

[0003] The increased number of cancer cases reported in the United States, and, indeed, around the world, is a major concern. Currently there are only a handful of treatments available for specific types of cancer, and these provide no guarantee of success. In order to be most effective, these treatments require not only an early detection of the malignancy, but a reliable assessment of the severity of the malignancy.

[0004] The incidence of breast cancer, a leading cause of death in women, has been gradually increasing in the United States over the last thirty years. Its cumulative risk is relatively high; 1 in 8 women are expected to develop some type of breast cancer by age 85 in the United States. In fact, breast cancer is the most common cancer in women and the second

most common cause of cancer death in the United States. In 1997, it was estimated that 181,000 new cases were reported in the U.S., and that 44,000 people would die of breast cancer (Parker *et al.*, *CA Cancer J. Clin.* 47:5-27 (1997); Chu *et al.*, *J. Nat. Cancer Inst.* 88:1571-1579 (1996)). While mechanism of tumorigenesis for most breast carcinomas is largely unknown, there are genetic factors that can predispose some women to developing breast cancer (Miki *et al.*, *Science*, 266:66-71(1994)).

[0005] Sporadic tumors, those not currently associated with a known germline mutation, constitute the majority of breast cancers. It is also likely that other, non-genetic factors also have a significant effect on the etiology of the disease. Regardless of the cancer's origin, breast cancer morbidity and mortality increases significantly if it is not detected early in its progression. Thus, considerable effort has focused on the early detection of cellular transformation and tumor formation in breast tissue.

[0006] A marker-based approach to tumor identification and characterization promises improved diagnostic and prognostic reliability. Typically, the diagnosis of breast cancer requires histopathological proof of the presence of the tumor. In addition to diagnosis, histopathological examinations also provide information about prognosis and selection of treatment regimens. Prognosis may also be established based upon clinical parameters such as tumor size, tumor grade, the age of the patient, and lymph node metastasis.

[0007] Diagnosis and/or prognosis may be determined to varying degrees of effectiveness by direct examination of the outside of the breast, or through mammography or other X-ray imaging methods (Jatoi, *Am. J. Surg.* 177:518-524 (1999)). The latter approach is not without considerable cost, however. Every time a mammogram is taken, the patient incurs a small risk of having a breast tumor induced by the ionizing properties of the radiation used during the test. In addition, the process is expensive and the subjective interpretations of a technician can lead to imprecision. For example, one study showed major clinical disagreements for about one-third of a set of mammograms that were interpreted individually by a surveyed group of radiologists. Moreover, many women find that undergoing a mammogram is a painful experience. Accordingly, the National Cancer Institute has not recommended mammograms for women under fifty years of age, since this group is not as likely to develop breast cancers as are older women. It is compelling to note, however, that while only about 22% of breast cancers occur in women under fifty, data suggests that breast cancer is more aggressive in pre-menopausal women.

[0008] In clinical practice, accurate diagnosis of various subtypes of breast cancer is important because treatment options, prognosis, and the likelihood of therapeutic response all

vary broadly depending on the diagnosis. Accurate prognosis, or determination of distant metastasis-free survival could allow the oncologist to tailor the administration of adjuvant chemotherapy, with women having poorer prognoses being given the most aggressive treatment. Furthermore, accurate prediction of poor prognosis would greatly impact clinical trials for new breast cancer therapies, because potential study patients could then be stratified according to prognosis. Trials could then be limited to patients having poor prognosis, in turn making it easier to discern if an experimental therapy is efficacious.

[0009] To date, no set of satisfactory predictors for prognosis based on the clinical information alone has been identified. Many have observed that the ER status has a dominant signature in the breast tumor gene expression profiling. See West *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 98:11462 (2001); van 't Veer *et al.*, *Nature* 415:530 (2002); Sorlie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:8418 (2003); Perou *et al.*, *Nature* 406:747 (2000); Gruvberger *et al.*, *Cancer Res.* 61:5979 (2001); Sotiriou *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:10393 (2003). It is generally accepted that there is some relationship between patient survival and ER status. van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002); Surowiak *et al.*, *Folia Histochem. Cytobiol.* 39:143 (2001); Pichon *et al.*, *Br. J. Cancer* 73:1545 (1996); Collett *et al.*, *J. Clin. Pathol.* 49:920 (1996). *BRCA1* mutations are related to the familial cancer susceptibility. Biesecker *et al.*, *JAMA* 269:1970 (1993); Easton *et al.*, *Cancer Surv.* 18:95 (1993). Age is also considered to be a prognosis factor since young cancer patients tend to have poor tumors. Maggard *et al.*, *J. Surg. Res.* 113:109 (2003). Lymph node status is a factor in deciding the treatment. Eifel *et al.*, *J. Natl. Cancer Inst.* 93:979 (2001).

[0010] The discovery and characterization of *BRCA1* and *BRCA2* has recently expanded our knowledge of genetic factors which can contribute to familial breast cancer. Germ-line mutations within these two loci are associated with a 50 to 85% lifetime risk of breast and/or ovarian cancer (Casey, *Curr. Opin. Oncol.* 9:88-93 (1997); Marcus *et al.*, *Cancer* 77:697-709 (1996)). Only about 5% to 10% of breast cancers, however, are associated with breast cancer susceptibility genes, *BRCA1* and *BRCA2*. The cumulative lifetime risk of breast cancer for women who carry the mutant *BRCA1* is predicted to be approximately 92%, while the cumulative lifetime risk for the non-carrier majority is estimated to be approximately 10%. *BRCA1* is a tumor suppressor gene that is involved in DNA repair and cell cycle control, which are both important for the maintenance of genomic stability. More than 90% of all mutations reported so far result in a premature truncation of the protein product with abnormal or abolished function. The histology of breast cancer in *BRCA1* mutation carriers differs from that in sporadic cases, but mutation analysis is the only way to find the carrier.

Like *BRCA1*, *BRCA2* is involved in the development of breast cancer, and like *BRCA1* plays a role in DNA repair. However, unlike *BRCA1*, it is not involved in ovarian cancer.

[0011] Other genes have been linked to breast cancer, for example c-erb-2 (*HER2*) and p53 (Beenken *et al.*, *Ann. Surg.* 233(5):630-638 (2001). Overexpression of c-erb-2 (*HER2*) and p53 have been correlated with poor prognosis (Rudolph *et al.*, *Hum. Pathol.* 32(3):311-319 (2001), as has been aberrant expression products of *mdm2* (Lukas *et al.*, *Cancer Res.* 61(7):3212-3219 (2001) and cyclin1 and p27 (Porter & Roberts, International Publication WO98/33450, published August 6, 1998).

[0012] The detection of *BRCA1* or *BRCA2* mutations represents a step towards the design of therapies to better control and prevent the appearance of these tumors. Recently, many studies have used gene expression profiling to analyze various cancers, and those studies have provided new diagnosis and prognosis information in the molecular level. See Zajchowski *et al.*, "Identification of Gene Expression Profiled that Predict the Aggressive Behavior of Breast Cancer Cells," *Cancer Res.* 61:5168 (2001); West *et al.*, "Predicting the Clinical Status of Human Breast Cancer by Using Gene Expression Profiles," *Proc. Natl. Acad. Sci. U.S.A.* 98:11462 (2001); van 't Veer *et al.*, "Gene Expression Profiling Predicts the Outcome of Breast Cancer," *Nature* 415:530 (2002); Roberts *et al.*, "Diagnosis and Prognosis of Breast Cancer Patients," WO 02/103320; Sorlie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:8418 (2003); Perou *et al.*, *Nature* 406:747 (2000); Khan *et al.*, *Cancer Res* 58, 5009 (1998); Golub *et al.*, *Science* 286, 531 (1999); DeRisi *et al.*, *Nat. Genet.* 14:457 (1996); Alizadeh *et al.*, *Nature* 403, 503 (2000). Methods for the identification of informative genesets for various cancers have also been described. See Roberts *et al.*, "Diagnosis and Prognosis of Breast Cancer Patients," WO 02/103320; Golub *et al.*, United States Patent No. 6,647,341.

[0013] Genesets have been identified that are informative for differentiating individuals having, or suspected of having, breast cancer based on estrogen receptor (ER) status, or *BRCA1* mutation vs. sporadic (*i.e.*, other than *BRCA1*-type) mutation status. See Roberts *et al.*, WO 02/103320; van't Veer *et al.*, *Nature* 415:530 (2001). Genesets have also been identified that enable the classification of sporadic tumor-type individuals as those who will likely have no metastases within five years of initial diagnosis (*i.e.*, individuals with a good prognosis) or those who will likely have a metastasis within five years of initial diagnosis (*i.e.*, those having a poor prognosis). Roberts, *supra*; van't Veer, *supra*.

[0014] Roberts *et al.* WO 02/103320 describes a 70-gene set, useful for the prognosis of breast cancer, which outperformed clinical measures of prognosis, and which showed good

potential in selecting good outcome patients, thereby avoiding over-treatment. van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002). The expression of genes with most predictive value, however, were not homogeneous among poor patients, suggesting the need for improvement.

[0015] Although the patterns of gene expression as described in Roberts *et al.* were correlated with existing clinical indicators such as estrogen receptor and *BRCAl* status, clinical measures were not incorporated. Furthermore, although the poor-outcome group in particular showed heterogeneity in expression pattern, the best classifier decision rule found during these studies was a fairly simple one based on the similarity of a patient profile to the average profile of a good-outcome training group.

[0016] It is evident that breast cancer is the result of more than one type of molecular event. Likewise, a variety of other conditions, such as other cancers; non-cancer diseases such as diabetes, autoimmune or neurodegenerative disorders, obesity; etc., are also the result of more than one molecular event. Moreover, an individual's response to exposure to particular environmental conditions, for example, exposure to natural or man-made agents, such as toxins, pollutants, drugs, food additives, etc., likely result from more than one molecular event. Thus, there exists a need for improved prognostic methods so that appropriate courses of prophylaxis and/or therapy may be provided. Genesets having improved prognostic power can be identified by first identifying discrete subsets of individuals based on genotypic or phenotypic characteristics relevant to the disease or condition, and then identifying genesets informative for prognosis within those subsets of patients. Individuals having the condition, or who are suspected of having the condition, such as breast cancer, would then be provided therapies appropriate to the molecular mechanisms underlying the condition. The present invention provides such methods for breast cancer, and for other cancers, diseases or conditions.

3. SUMMARY OF THE INVENTION

[0017] The present invention provides methods of identifying relevant subsets of conditions, and the identification of markers relevant to those subsets, for example, for prognosis of individuals classifiable into one of those subsets. The invention further provides sets of markers useful for the prognosis of individuals having breast cancer, wherein those patients have been classified according to one or more characteristics of breast cancer.

[0018] Thus, the present invention provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) classifying each of a plurality of samples or individuals on the

basis of one or more phenotypic or genotypic characteristics of said condition into a plurality of first classes; and (b) identifying within each of said first classes a first set of genes or markers informative for said condition, wherein said first set of genes or markers within each of said first classes is unique to said class relative to other first classes. In a specific embodiment, this method further comprises additionally classifying into a plurality of second classes said samples or individuals in at least one of said first classes on the basis of a phenotypic or genotypic characteristic different than that used in said classifying step (a); and identifying within at least one of said second classes a second set of informative genes or markers, wherein said second set of informative genes or markers within each of said second classes is unique to said second class relative to other first and second classes.

[0019] The invention further provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics into a plurality of first classes; (b) classifying at least one of said first classes into a plurality of second classes on the basis of phenotypic or genotypic characteristic different than that used in said classifying step (a); and (c) identifying within at least one of said first classes or said second classes a set of genes or markers informative for said condition, wherein said second set of genes or markers is unique to said class relative to other first and second classes.

[0020] The invention further provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) selecting a first characteristic from said plurality of phenotypic or genotypic characteristics; (b) identifying at least two first condition classes differentiable by said first characteristic; (c) selecting a plurality of individuals classifiable into at least one of said first condition classes; and (d) identifying in samples derived from each of said plurality of individuals a set of genes or markers informative for said condition within said at least one of said first condition classes.

[0021] The invention further provides a method of classifying an individual with a condition as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual into one of a plurality of patient classes, said patient classes being differentiated by one or more phenotypic, genotypic or clinical characteristics of said condition; (b) determining the level of expression of a plurality of genes or their encoded proteins in a cell sample taken from the individual relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins informative for prognosis of the patient

class into which said individual is classified; and (c) classifying said individual as having a good prognosis or a poor prognosis on the basis of said level of expression. In a specific embodiment, said condition is cancer, said good prognosis is the non-occurrence of metastases within five years of initial diagnosis, and said poor prognosis is the occurrence of metastases within five years of initial diagnosis. In a more specific embodiment, said cancer is breast cancer. In another specific embodiment, said control is the average level of expression of each of said plurality of genes or their encoded proteins across a plurality of samples derived from individuals identified as having a poor prognosis. In a more specific embodiment, said classifying step (c) is carried out by a method comprising comparing the level of expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a poor prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in said control more strongly than would be expected by chance. In another specific embodiment, said control is the average level of expression of each of said plurality of genes or their encoded proteins across a plurality of samples derived from individuals identified as having a good prognosis. In a more specific embodiment, said classifying in step (c) is carried out by a method comprising comparing the level expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a good prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in said control more strongly than would be expected by chance. In another specific embodiment, said plurality of patient classes comprises ER^{-} , *BRCAI* individuals; ER^{-} , sporadic individuals; ER^{+} , ER/AGE high individuals; ER^{+} , ER/AGE low, LN^{+} individuals; and ER^{+} , ER/AGE low, LN^{-} individuals.

[0022] The invention further provides a method of classifying a breast cancer patient as having a good prognosis or a poor prognosis comprising: (a) classifying said breast cancer patient as ER^{-} , *BRCAI*; ER^{-} , sporadic; ER^{+} , ER/AGE high; ER^{+} , ER/AGE low, LN^{+} ; or ER^{+} , ER/AGE low, LN^{-} ; (b) determining the level of expression of a first plurality of genes in a cell sample taken from said breast cancer patient relative to a control, said first plurality of genes comprising two of the genes corresponding to the markers in Table 1 if said breast cancer patient is classified as ER^{-} , *BRCAI*; in Table 2 if said breast cancer patient is classified as ER^{-} sporadic; in Table 3 if said breast cancer patient is classified as ER^{+} , ER/AGE high; in Table 4 if said breast cancer patient is classified as ER^{+} , ER/AGE low,

LN⁺; or in Table 5 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁻; and (c) classifying said breast cancer patient as having a good prognosis or a poor prognosis on the basis of the level of expression of said first plurality of genes, wherein said breast cancer patient is “ER/AGE high” if the ratio of the log₁₀(ratio) of ER gene expression to age exceeds a predetermined value, and “ER/AGE low” if the ratio of the log₁₀(ratio) of ER gene expression to age does not exceed said predetermined value. In a specific embodiment, said control is the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁻, *BRCA1* individuals, if said breast cancer patient is ER⁻, *BRCA1*; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁻, sporadic individuals if said breast cancer patient is ER⁻, sporadic; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE high individuals, if said breast cancer patient is ER⁺, ER/AGE high; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE low, LN⁺ individuals where said breast cancer patient is ER⁺, ER/AGE low, LN⁺; or the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE low, LN⁻ individuals where said breast cancer patient is ER⁺, ER/AGE low, LN⁻. In a more specific embodiment, each of said individuals has a poor prognosis. In another more specific embodiment, each of said individuals has a good prognosis. In an even more specific embodiment, said classifying step (c) is carried out by a method comprising comparing the level of expression of each of said plurality of genes or their encoded proteins in a sample from said breast cancer patient to said control, and classifying said breast cancer patient as having a poor prognosis if said level of expression correlates with said average level of expression of the corresponding genes or their encoded proteins in said control more strongly than would be expected by chance. In another specific embodiment, said predetermined value of ER is calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said individual. In another specific embodiment, said individual is ER⁻, *BRCA1*, and said plurality of genes comprises two of the genes for which markers are listed in Table 1. In another specific embodiment, said individual is ER⁻, *BRCA1*, and said plurality of genes comprises all of the genes for which markers are listed in Table 1. In another specific embodiment, said individual is ER⁻, sporadic, and said plurality of genes comprises two of the genes for which markers are listed in Table 2. said individual is ER⁻, sporadic, and said plurality of genes comprises all of the genes for which markers are listed in Table 2. In another specific embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises two of the genes for which markers are listed in Table 3. said

individual is ER+, ER/AGE high, and said plurality of genes comprises all of the genes for which markers are listed in Table 3. In another specific embodiment, said individual is ER+, ER/AGE low, LN+, and said plurality of genes comprises two of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER+, ER/AGE low, LN+, and said plurality of genes comprises all of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER+, ER/AGE low, LN⁻, and said plurality of genes comprises two of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER+, ER/AGE low, LN⁻, and said plurality of genes comprises all of the genes for which markers are listed in Table 4. In another specific embodiment, the method further comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

[0023] In another embodiment, the invention provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising: (a) classifying said individual as ER⁻, *BRCA1*, ER⁻, sporadic; ER+, ER/AGE high; ER+, ER/AGE low, LN+; or ER+, ER/AGE low, LN⁻; (b) determining for said individual the level of expression of at least two genes for which markers are listed in Table 1 if said individual is classified as ER⁻, *BRCA1*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER+, ER/AGE high; Table 4 if said individual is classified as ER+, ER/AGE low, LN+; or Table 5 if said individual is classified as ER+, ER/AGE low, LN⁻; (c) determining whether said individual has a pattern of expression of said at least two genes that correlates with a good prognosis or a poor prognosis; and (d) assigning said individual to one category in a clinical trial if said individual has a good prognosis, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis. In a specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a). In another specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of any other clinical, phenotypic or genotypic characteristic of breast cancer. In another specific embodiment, said method further comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis of breast cancer, and determining from the expression of said second plurality of genes, in addition to said first plurality of genes, whether said individual has a good prognosis or a poor prognosis.

[0024] The invention further provides a microarray comprising probes complementary and hybridizable to a plurality of the genes for which markers are listed in any of Tables 1-5. The invention further provides a microarray comprising probes complementary and hybridizable to a plurality of the genes for which markers are listed in Table 1, each of the genes for which markers are listed in Table 1, a plurality of the genes for which markers are listed in Table 2, each of the genes for which markers are listed in Table 2, a plurality of the genes for which markers are listed in Table 3, each of the genes for which markers are listed in Table 3, a plurality of the genes for which markers are listed in Table 4, each of the genes for which markers are listed in Table 4, a plurality of the genes for which markers are listed in Table 5, or each of the genes for which markers are listed in Table 5. The invention further provides any one of the above microarrays, wherein said probes are at least 50% of the probes on said microarray. The invention further provides any one of the above microarrays, wherein said probes are at least 90% of the probes on said microarray. The invention further provides microarray comprising probes complementary and hybridizable to a plurality of the genes for which markers are listed in any of Tables 1-5, wherein said probes are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 1; are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 2; are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 3; are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 4; and are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 5, wherein said probes, in total, are at least 50% of the probes on said microarray.

[0025] The invention further comprises a kit comprising any one of the above microarrays in a sealed container.

[0026] The invention further provides a method of identifying a set of genes informative for a condition, said condition having a plurality of phenotypic or genotypic characteristics such that samples may be categorized by at least one of said phenotypic or genotypic characteristics into at least one characteristic class, said method comprising: (a) selecting a plurality of samples from individuals having said condition; (b) identifying a first set of genes informative for said characteristic class using said plurality of samples; (c) predicting the characteristic class of each of said plurality of samples; (d) discarding samples for which said characteristic class is incorrectly predicted; (e) repeating steps (c) and (d) at least once; and (f) identifying a second set of genes informative for said characteristic class using samples in said plurality of samples remaining after step (e).

[0027] The invention further provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising: (a) classifying the individual into one of a plurality of condition categories differentiated by at least one genotypic or phenotypic characteristic of the condition; (b) determining the level of expression, in a sample derived from said individual, of a plurality of genes informative for said condition category; (c) determining whether said level of expression of said plurality of genes indicates that the individual has a good prognosis or a poor prognosis; and (d) assigning the individual to a category in a clinical trial on the basis of prognosis.

[0028] The invention also provides a method for identifying one or more sets of informative genes or markers for a condition in an organism, comprising: (a) subdividing a plurality of individuals or samples derived therefrom of the organism subject to the condition into a plurality of classes based on one or more clinical, phenotypic or genotypic characteristics of the organism, wherein each class consists of a plurality of individuals or samples derived therefrom of the organism each of which having one or more clinical, phenotypic or genotypic characteristics specific for the class; and (b) attempting to identify for each of one or more of said plurality of classes a set of genes or markers informative for said condition in individuals in said class, wherein, if a set of genes or markers informative for said condition in individuals in said class is obtained for any of said one or more of said plurality of classes, said set of genes or markers is taken as a set of informative genes or markers for said condition in said organism.

[0029] In one embodiment, the method further comprises, for each of one or more of said classes in which a set of genes or markers informative for said condition in individuals in said class cannot be obtained, repeating said steps (a) and (b) on said plurality of individuals or samples derived therefrom in said class such that said plurality of individuals or samples derived therefrom in said class is subdivided into a plurality of additional classes based on one or more clinical, phenotypic or genotypic characteristics of said organism which are different from those used for defining said class, wherein for each of said plurality of additional classes, if a set of genes or markers informative for said condition in individuals in said class is obtained, said set of genes or markers is taken as a set of informative genes or markers for said condition in said organism.

[0030] The invention also provides a method for identifying one or more sets of informative genes or markers for a condition in an organism, comprising: (a) subdividing a plurality of individuals or samples derived therefrom of said organism subject to said condition into a plurality of classes based on one or more clinical, phenotypic or genotypic characteristics of

said organism, wherein each said class consists of a plurality of individuals or samples derived therefrom of said organism each having said one or more clinical, phenotypic or genotypic characteristics specific for said class; (b) attempting to identify for each of one or more of said plurality of classes a set of genes or markers informative for said condition in individuals in said class, wherein if a set of genes or markers informative for said condition in individuals in said class is identified for any of said one or more of said classes, said set of genes or markers is taken as a set of informative genes or markers for a condition in said organism; and (c) for each of one or more of said classes in which a set of genes or markers informative for said condition in individuals in said class cannot be obtained, repeating said steps (a) and (b) on said plurality of individuals or samples derived therefrom in said class such that said plurality of samples or individuals in said class is subdivided into a plurality of additional classes based on one or more clinical, phenotypic or genotypic characteristics of said organism which are different from those used for defining said class, wherein for each of one or more of said plurality of additional classes, if a set of genes or markers informative for said condition in individuals in said class is obtained, said set of genes or markers is taken as a set of informative genes or markers for a condition in said organism.

[0031] In the methods of the invention, the condition can be a type of cancer. In such an embodiment, each of said sets of genes or markers can be informative of prognosis of individuals in a corresponding class. In one embodiment, the condition is breast cancer, and the one or more clinical, phenotypic or genotypic characteristics comprise age, ER level, ER/AGE, BRAC1 status, and lymph node status.

[0032] In one embodiment, the methods of the invention further comprise generating a template profile comprising measurements of levels of genes or markers of the set of informative genes or markers for said class representative of levels of the genes or markers in a plurality of patients having a chosen prognosis level.

[0033] The invention also provides a method for predicting a breast cancer patient as having a good prognosis or a poor prognosis, comprising: (a) classifying said breast cancer patient into one of the following classes: (a1) ER⁻, *BRCA1*; (a2) ER⁻, sporadic; (a3) ER⁺, ER/AGE high; (a4) ER⁺, ER/AGE low, LN⁺; or (a5) ER⁺, ER/AGE low, LN⁻; (b) determining a profile comprising measurements of a plurality of genes or markers in a cell sample taken from said breast cancer patient, said plurality of genes markers comprising at least two of the genes or markers corresponding to the markers in (b1) Table 1 if said breast cancer patient is classified as ER⁻, *BRCA1*; (b2) Table 2 if said breast cancer patient is classified as ER⁻ sporadic; (b3) Table 3 if said breast cancer patient is classified as ER⁺, ER/AGE high; (b4)

Table 4 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁺; or (b5) Table 5 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁻; and (c) classifying said breast cancer patient as having a good prognosis or a poor prognosis based on said profile of said plurality of genes or markers, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, wherein said ER/AGE is a metric of said ER level relative to the age of said patient, and wherein LN⁺ designates a greater than 0 lymph nodes status in said patient and LN⁻ designates a 0 lymph nodes status in said patient.

[0034] In one embodiment, step (c) is carried out by a method comprising comparing said profile to a good prognosis template and/or a poor prognosis template, and wherein said patient is classified as having a good prognosis if said profile has a high similarity to a good prognosis template or has a low similarity to a poor prognosis template or as having a poor prognosis if said profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template. A good prognosis template comprises measurements of said plurality of genes or markers representative of levels of said genes or markers in a plurality of good outcome patients, while a poor prognosis template comprises measurements of said plurality of genes or markers representative of levels of said genes or markers in a plurality of poor outcome patients. Here a good outcome patient is a breast cancer patient who has non-reoccurrence of metastases within a first period of time after initial diagnosis, while a poor outcome patient is a patient who has reoccurrence of metastases within a second period of time after initial diagnosis.

[0035] In another embodiment, the methods for predicting the prognosis of a breast cancer patient further comprise determining said profile, said ER level, said LN status, and/or, said ER/AGE. In one embodiment, said profile is an expression profile comprising measurements of a plurality of transcripts in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

[0036] In one embodiment, said expression profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample.

[0037] In one embodiment, the measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

[0038] In one embodiment, the similarity of said expression profile to said good or poor prognosis template is represented by a correlation coefficient between said expression profile and said good or poor prognosis template, respectively, and a correlation coefficient greater than a correlation threshold, e.g., 0.5, indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

[0039] In another embodiment, the similarity of said expression profile to said good or poor prognosis template is represented by a distance between said cellular constituent profile and said good or poor prognosis template, respectively, and a distance less than a given value indicates a high similarity and said distance equal to or greater than said given value indicates a low similarity.

[0040] In another embodiment, said profile comprises measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome patients.

[0041] In one embodiment, said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor, e.g., the estrogen receptor α gene, in said patient relative to expression level of said gene in said control sample, and said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65.

[0042] In one embodiment, said ER/AGE is classified as high if said ER level is greater than $c \cdot (\text{AGE} - d)$, and said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

[0043] In a specific embodiment, said estrogen receptor level is measured by a polynucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, said control sample is a pool of breast cancer cells of different patients, and $c = 0.1$ and $d = 42.5$.

[0044] In one embodiment, said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients. In another embodiment, said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript of said gene encoding said estrogen receptor.

[0045] In one embodiment, said individual is ER⁻, *BRCAl*, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 1. In one embodiment, said individual is ER⁻, *BRCAl*, and said plurality of genes comprises all of the genes for which markers are listed in Table 1.

[0046] In another embodiment, the individual is ER⁻, sporadic, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 2. In one embodiment, said individual is ER⁻, sporadic, and said plurality of genes comprises all of the genes for which markers are listed in Table 2.

[0047] In still another embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 3. In one embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises all of the genes for which markers are listed in Table 3.

[0048] In still another embodiment, said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 4. In one embodiment, said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

[0049] In still another embodiment, said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 4. In one embodiment, the individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

[0050] In one embodiment, said profile further comprises one or more genes for which markers are not found in Tables 1-5, which are informative for prognosis.

[0051] The invention also provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising assigning said individual to one category in a clinical trial if said individual has a good prognosis as determined by any one of the methods described above, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis as determined by any one of the methods described above.

[0052] In one embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual based on said profile, said ER level, said LN status, and/or, said ER/AGE.

[0053] In one embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of one or more other clinical, phenotypic or genotypic characteristic of breast cancer.

[0054] In one embodiment, the method further comprises determining in said cell sample the levels of expression of said one or more genes for which markers are not found in Tables 1-5, and determining from said expression levels of said one or more genes, whether said individual has a good prognosis or a poor prognosis.

4. BRIEF DESCRIPTION OF THE DRAWINGS

[0055] FIG. 1 depicts the decision tree that resulted in the five patient subsets used to identify informative prognosis-related genes.

[0056] FIG. 2: Relationship between ER level and age. (A) Scatter plot of ER vs. age for ER+ patients. Black dots indicate metastases free samples, and gray dots indicate metastases samples. It appears that patients of ER+ group can be subdivided into “ER+, ER/AGE high” group (above the black line) and “ER+, ER/AGE low” (below the black line) group. The black line is approximated by $ER = 0.1 * (AGE - 42.5)$, and the dashed line by $ER = 0.1 * (age - 50)$. Within each population, the ER level also increases with age. (B) Age distribution of all patients in ER+ samples. A bimodal distribution is observed. (C) ER-modulated age (age – 10*) distribution of all patients in ER+ samples. A bimodal distribution is observed. (D) Age distribution of samples with metastasis. (E) ER-modulated age distribution of samples with metastasis. The three peaks appearing in this distribution suggest a polymorphism.

[0057] FIG. 3. Performance of classifier for the “ER⁻/sporadic” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0058] FIG. 4. Performance of classifier for the “ER+, ER/AGE high” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, and squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0059] FIG. 5. Performance of classifier for the “ER+, ER/AGE low/LN⁻” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, and squares indicates samples with metastases. Dashed line indicates the threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0060] FIG. 6. Performance of classifier for the “ER+, ER/AGE low/LN⁺” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0061] FIG. 7. Performance of classifier for the “ER⁻, *BRCAl*” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group

(poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0062] FIG. 8. Heatmaps of genes representing key biological functions in subgroups of patients: A: Cell cycle genes are predictive of outcome in patients with ER/age high. B: Cell cycle genes are not predictive of outcome in “ER- and sporadic” patients C: Glycolysis genes are predictive of outcome in patients with ER/age low and LN-. D: Glycolysis genes are not predictive of outcome in ‘ER- & BRCA1” patients.

5. DETAILED DESCRIPTION OF THE INVENTION

5.1 INTRODUCTION

[0063] The present invention provides methods for classifying individuals having a condition, such as a disease, into one or more subsets of individuals, where individuals in each subset are characterized by one or more phenotypic or genotypic characteristics of the condition. The individuals may be eukaryotes or prokaryotes, may be animals such as mammals, including but not limited to humans, primates, rodents, felines, canines, etc., birds, reptiles, fish, etc. “Individuals” as used herein also encompasses single-celled organisms, or colonies thereof, such as bacteria and yeast. The condition may be a disease, such as cancer, and may be a specific cancer, such as breast cancer. The condition may also be an environmental condition, such as exposure to a toxin, pollutant, drug, proximity to urban or industrial areas, etc.

[0064] The present invention provides methods of determining the prognosis of individuals having a condition, such as cancer, for example, breast cancer, or who are suspected of having the condition, by the use of a combination of clinical, biological or biochemical parameters of the condition and gene expression pattern data. For prognosis, the parameters selected preferably relate to or affect the progression and/or outcome of the condition. The pattern of gene expression within a subset of individuals having the particular condition leads to the identification of sets of genes within a subset that is informative for that subset, for example, for prognosis within that subset. In general, the successful identification of sets of genes informative for prognosis within a particular subset justifies the selection of the plurality of clinical, biological or biochemical parameters of the condition on which division of individuals into condition subsets is based.

[0065] In the example of breast cancer, patient groups are first classified according to at least one of age, lymph node (LN) status, estrogen receptor (ER) level, and *BRCA1* mutation status into discrete patient subsets. These clinical factors have been implicated in tumor

etiology as well as differences in disease outcome. These characteristics are not limiting; other genotypic or phenotypic characteristics of breast cancer, for example, tumor grade, tumor size, tumor cell type, etc., may also be used, alone or in combination with those listed herein, in order to classify individuals. The differences in gene expression or in tumor fate related to these parameters likely represent differences in tumor origin and tumor genesis, and are therefore good candidates for tumor stratification. Genesets informative for prognosis within each subset are then identified. New breast cancer patients are then classified using the same criteria, and a prognosis is made based on the geneset specific for the patient subset into which the patient falls. In the process of constructing a prognosis classifier within each patient subset, particular attention is paid to the homogeneous patterns related to the tumor outcome. Emergence of such homogeneous prognosis patterns may indicate the most common mechanism to metastasis within a subset. At the same time, successful identification of such patterns also justifies the parameters being used for the tumor stratification. To differentiate this approach from an mRNA-alone approach, the current approach of integrating clinical data with the gene expression data is referred to herein as a “comprehensive prognosis”.

5.2 DEFINITIONS

[0066] As used herein, “*BRCA1* tumor” or “*BRCA1* type” means a tumor having cells containing a mutation of the *BRCA1* locus.

[0067] The “absolute amplitude” of correlation means the absolute value of the correlation; e.g., both correlation coefficients -0.35 and 0.35 have an absolute amplitude of 0.35.

[0068] “Marker” means a cellular constituent, or a modification of a cellular constituent (e.g., an entire gene, EST derived from that gene, a protein encoded by that gene, post-translational modification of the protein, etc.) the expression or level of which changes between certain conditions. Where a change in a characteristic of the constituent correlates with a certain condition, the constituent is a marker for that condition.

[0069] “Marker-derived polynucleotides” means the RNA transcribed from a marker gene, any cDNA or cRNA produced therefrom, and any nucleic acid derived therefrom, such as synthetic nucleic acid having a sequence derived from the gene corresponding to the marker gene.

[0070] A “similarity value” is a number that represents the degree of similarity between two things being compared. For example, a similarity value may be a number that indicates the overall similarity between a patient’s expression profile of specific phenotype-related

markers and a template specific to that phenotype (for instance, the similarity to a “good prognosis” template, where the phenotype is a good prognosis). The similarity value may be expressed as a similarity metric, such as a correlation coefficient, or may simply be expressed as the expression level difference, or the aggregate of the expression level differences, between a patient sample and a template.

[0071] A “patient subset” is a group of individuals, all of whom have a particular condition, or are subject to a particular condition, which is distinguished from other individuals having that condition by one or more phenotypic, genotypic or clinical characteristics of the condition, or of a response to the condition. For example, where the condition is breast cancer, individuals may belong to an “ER⁺” or an “ER⁻” patient subset, or may belong to a particular age group patient subset.

[0072] A gene and/or marker is “informative” for a condition, phenotype, genotype or clinical characteristic if the expression of the gene or marker is correlated or anticorrelated with the condition, phenotype, genotype or clinical characteristic to a greater degree than would be expected by chance.

[0073] An individual of a given age can be classified as “ER/AGE high” if the individual’s ER level is higher than a threshold value for the given age. The threshold can be age-dependent, i.e., a different threshold for each different age. In one embodiment, the age-dependent threshold value is calculated as $c \cdot (AGE - d)$, where c is a coefficient, AGE is the age of the patient, and d is an age threshold. The parameters c and d depend on the ER level and AGE used. They can be determined by fitting patients’ ER level-age distribution to a bimodal distribution of two subgroups each having a different ER level-age dependence. In a specific embodiment, $c = 0.1$ and $d = 42.5$ is used for ER levels represented by a log(ratio) of ER expression level. Thus, for example, the threshold for a 45-year old individual in this embodiment is $0.1 (45 - 42.5)$, or 0.25, and if the log(ratio) of ER expression level of the individual is equal to or greater than 0.25, the individual is classified as “ER/AGE high”; otherwise, the individual is classified as “ER/AGE low.”

5.3 IDENTIFICATION OF DIAGNOSTIC AND PROGNOSTIC MARKER SETS

5.3.1 IDENTIFICATION OF CONDITION SUBSETS

[0074] The present invention provides methods of identifying sets of genes and/or markers useful in the diagnosis and prognosis of breast cancer. More generally, the invention also provides methods of identifying sets of genes and/or markers useful in the diagnosis or prognosis of other cancers, and even more generally, of identifying sets of genes and/or

markers useful in the differentiation between subgroups of individuals having a particular condition, such as a disease or exposure to a particular environmental condition.

[0075] The method may be applied to any condition for which a plurality of phenotypic or genotypic subsets may be identified. The condition may be a disease; for example, the condition may be cancer, an autoimmune disease, an inflammatory disease, an infectious disease, a neurological disease, a degenerative disease, etc. The condition may be environmental; for example, the condition may be a particular diet, geographic location, etc.; the condition may be exposure to a compound, including, for example, a drug, a toxin, a carcinogen, a foodstuff, a poison, an inhaled compound, an ingested compound, etc.; the condition may be a particular genetic background or predisposition to a medical condition; etc.

[0076] Where the condition is cancer, the condition may be any cancer, for example, without limitation: leukemias, including acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic leukemia, promyelocytic leukemia, myelomonocytic leukemia, monocytic leukemia, and erythroleukemia; chronic leukemia, such as chronic myelocytic (granulocytic) leukemia or chronic lymphocytic leukemia; polycythemia vera; lymphomas, such as Hodgkin's disease and non-Hodgkin's disease; multiple myeloma; Waldenström's macroglobulinemia; heavy chain disease; solid tumors, such as sarcomas and carcinomas, fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilms' tumor, cervical cancer, testicular tumor, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, menangioma, melanoma, neuroblastoma, or retinoblastoma; etc.

[0077] Rather than stratifying individuals, such as patients or tumor samples derived from patients, by gene expression patterns in the first instance, the method of identifying sets of genes informative for a condition begins by identifying phenotypic, genotypic or clinical

subsets of individuals within the larger class of individuals having or affected by the condition.

[0078] In one embodiment, the condition is cancer, and the subsets are distinguished by phenotypic, genotypic, and/or clinical characteristics of the cancer. In this embodiment, groups of individuals are classified according to one or more phenotypic, genotypic, or clinical characteristics relevant to the cancer into patient subsets. At any step in the process of subdividing a patient population into patient subsets, the expression level of one or more genes may be determined in order to identify whether a prognosis-informative set of genes may be identified for the particular patient subset. If an informative gene set is identified, but is not as informative as desired, the patient subset may be further divided and a new geneset identified. These subsets may be further subdivided. For example, a group of individuals affected by a particular cancer may be classified first on the basis of a phenotypic, genotypic or clinical characteristic A into subsets S1 and S2. The levels of expression of a plurality of genes are then determined in tumor samples taken from individuals that fall within subsets S1 or S2 in order to identify sets of genes informative for prognosis within these subsets. Subsets S1 and S2 may then each be subdivided into two or more subsets based on other phenotypic, genotypic or clinical characteristics. The basis for subdivision, if performed, need not be the same for S1 and S2. For example, in various embodiments, S1 is not subdivided, while S2 is subdivided on the basis of characteristic B; or S1 is subdivided based on characteristic B while S2 is not subdivided; or S1 and S2 are both subdivided on the basis of characteristic B; or S1 is subdivided based on characteristic B, while S2 is subdivided according to characteristic C; and so on. For a particular decision matrix leading to a plurality of patient subsets, the preferred outcome is a prognosis-informative set of genes for each patient subset. Different decision matrices may lead to different patient subsets, which, in turn, may result in different sets of prognosis-informative genes.

[0079] In the specific example of breast cancer, a plurality of phenotypic, genotypic or clinical indications are used to classify a patient as being a member of one of a plurality of patient subsets, wherein the indications are medically, biochemically or genetically relevant to breast cancer. For example, a group of patients may be classified into patient subsets based on criteria including, but not limited to, estrogen receptor (ER) status, type of tumor (*i.e.*, *BRCA1*-type or sporadic), lymph node status, grade of cancer, invasiveness of the tumor, or age. "BRCA1-type" indicates that the *BRCA1* mutation is present. In each classification step, a group of cancer patients may be classified into only two classes, for example, ER+ or ER⁻, or into three or more subsets (for example, by tumor grade), depending upon the

characteristic used to determine the subsets. As used herein, “ER+” indicates that the estrogen receptor is expressed at some elevated level; for example, it may indicate that the estrogen receptor is detectably expressed, or may indicate that more than 10% of cells are histologically stained for the receptor, etc. Conversely, “ER–” indicates that the estrogen receptor is expressed at a reduced level or not at all; for example, it may indicate that the receptor is not detectably expressed, or that 10% or less of cells are histologically stained for the receptor, etc. Marker gene sets optimized for each phenotypic class are preferably determined after the subsets are established. Where informative markers for a particular patient subset, distinguished from another subset by a particular characteristic of the condition of interest, cannot be determined, the subset may be further divided by another characteristic of the condition to create a plurality of second patient subsets, whereupon genes informative for these second patient subsets may be identified.

[0080] FIG. 1 depicts the process, described in the Examples, of subdivision of a collection of breast cancer patients according to phenotypic and genotypic characteristics relevant to breast cancer, in preparation for identification of genes informative for prognosis. A collection of breast cancer tumor samples was first subdivided by estrogen receptor status. ER status was chosen because the presence or absence of the estrogen receptor greatly influences the expression of other genes. In the ER+ patient subset, it was noted that patients appeared to be bimodally distributed by ER level vs. age; that is, ER level dependence upon age tended to fall within two classes, as separated by the solid line in FIG. 2A. This bimodality was used to further subdivide ER+ individuals into “ER+, ER/AGE high” individuals and “ER+, ER/AGE low” individuals. A set of informative genes was identified for the ER+, ER/AGE high patient subset. An informative set was not identified for the ER+, ER/AGE low subset, however, so the subset of patients was further divided into LN+ and LN– individuals. Thus, in one embodiment, the present invention provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising (a) classifying each of a plurality of samples or individuals on the basis of one phenotypic or genotypic characteristic into a plurality of first classes; and (b) identifying within each of said first classes a set of informative genes or markers, wherein said set of informative genes or markers within each said first classes is unique to said class.

5.3.2 IDENTIFICATION OF MARKER SETS INFORMATIVE FOR PATIENT SUBSETS

[0081] Once a patient subset is identified, markers, such as genes, informative for a particular outcome, such as prognosis, may be identified. In one embodiment, the method for identifying marker sets is as follows. This example describes the use of genes and gene-derived nucleic acids as markers; however, proteins or other cellular constituents may be used as markers of the condition.

[0082] After extraction and labeling of target polynucleotides, the expression of a plurality of markers, such as genes, in a sample X is compared to the expression of the plurality markers in a standard or control. In one embodiment, the standard or control comprises target markers, such as polynucleotide molecules, derived from one or more samples from a plurality of normal individuals, or a plurality of individuals not exposed to a particular condition. For example, the control, or normal, individuals may be persons without the particular disease or condition of interest (*e.g.*, individuals not afflicted with breast cancer, where breast cancer is the disease of interest), or may be an individual not exposed to a particular environmental condition. The standard or control may also comprise target polynucleotide molecules, derived from one or more samples derived from individuals having a different form or stage of the same disease; a different disease or different condition, or individuals exposed or subjected to a different condition, than the individual from which sample X was obtained. The control may be a sample, or set of samples, taken from the individual at an earlier time, for example, to assess the progression of a condition, or the response to a course of therapy.

[0083] In a preferred embodiment, the standard or control is a pool of target polynucleotide molecules. However, where protein levels, or the levels of any other relevant biomolecule, are to be compared, the pool may be a pool of proteins or the relevant biomolecule. In a preferred embodiment in the context of breast cancer, the pool comprises samples taken from a number of individuals having sporadic-type tumors.

[0084] In another preferred embodiment, the pool comprises an artificially-generated population of nucleic acids designed to approximate the level of nucleic acid derived from each marker found in a pool of marker-derived nucleic acids derived from tumor samples. In another embodiment, the pool, also called a “mathematical sample pool,” is represented by a set of expression values, rather than a set of physical polynucleotides; the level of expression of relevant markers in a sample from an individual with a condition, such as a disease, is compared to values representing control levels of expression for the same markers in the mathematical sample pool. Such a control may be a set of values stored on a computer. Such artificial or mathematical controls may be constructed for any condition of interest.

[0085] In another embodiment specific to breast cancer, the pool is derived from normal or breast cancer cell lines or cell line samples. In a preferred embodiment, the pool comprises samples taken from individuals within a specific patient subset, *e.g.*, “ER+, ER/AGE high” individuals, wherein each of said individuals has a good prognosis, or each of said individuals has a poor prognosis. Of course, where, for example, expressed proteins are used as markers, the proteins are obtained from the individual’s sample, and the standard or control could be a pool of proteins from a number of normal individuals, or from a number of individuals having a particular state of a condition, such as a pool of samples from individuals having a particular prognosis of breast cancer.

[0086] The comparison may be accomplished by any means known in the art. For example, expression levels of various markers may be assessed by separation of target polynucleotide molecules (*e.g.*, RNA or cDNA) derived from the markers in agarose or polyacrylamide gels, followed by hybridization with marker-specific oligonucleotide probes. Alternatively, the comparison may be accomplished by the labeling of target polynucleotide molecules followed by separation on a sequencing gel. Polynucleotide samples are placed on the gel such that patient and control or standard polynucleotides are in adjacent lanes. Comparison of expression levels is accomplished visually or by means of densitometer. In a preferred embodiment, the expression of all markers is assessed simultaneously by hybridization to a microarray. In each approach, markers meeting certain criteria are identified as informative for the prognosis of breast cancer.

[0087] Marker genes are selected based upon significant difference of expression in a condition, such as a disease, as compared to a standard or control condition. Marker genes may be screened, for example, by determining whether they show significant variation within a set of samples of interest. Genes that do not show a significant amount of variation within the set of samples are presumed not to be informative for the disease or condition, and are not selected as markers for the disease or condition. Genes showing significant variation within the sample set are candidate informative genes for the disease or condition. The degree of variation may be estimated by calculating the difference of the expression of the gene, or ratio of expression between sample and control, within the set of samples. The expression, or ratio of expressions, may be transformed by any means, *e.g.*, linear or log transformation. Selection may be made based upon either significant up- or down regulation of the marker in the patient sample. Selection may also be made by calculation of the statistical significance (*i.e.*, the p-value) of the correlation between the expression of the marker and the disease and condition. Preferably, both selection criteria are used. Thus, in one embodiment of the

present invention, markers associated with prognosis of breast cancer within a patient subset are selected where the markers show both more than two-fold change (increase or decrease) in expression as compared to a standard, and the p-value for the correlation between the existence of breast cancer and the change in marker expression is no more than 0.01 (*i.e.*, is statistically significant).

[0088] In the context of the present invention, “good prognosis” indicates a desired outcome for a particular condition, especially a particular disease, and “poor prognosis” indicates an undesired outcome of the condition. For example, where the condition is cancer, a “good prognosis” may mean partial or complete remission, and “poor prognosis” may mean reappearance of the cancer after treatment. What constitutes “good prognosis” and “poor prognosis” is specific to the condition of interest, for example, specific to the particular cancer an individual suffers. For example, “good prognosis” for pancreatic cancer may be survival for one or two years after initial diagnosis, while “good prognosis” for Hodgkin’s disease may be survival for five years or more. In the specific example of breast cancer, “good prognosis” means the likelihood of non-reoccurrence of metastases within a period of 1, 2, 3, 4, 5 or more years after initial diagnosis, and “poor prognosis” means the likelihood of reoccurrence of metastasis within that period. In a more specific example, “good prognosis” means the likelihood of non-reoccurrence of metastases within 5 years after initial diagnosis, and “poor prognosis” means the likelihood of reoccurrence of metastasis within that period.

[0089] In a more specific embodiment for cancer, for example, breast cancer, using a number of breast cancer tumor samples, markers are identified by calculation of correlation coefficients ρ between the clinical category or clinical parameter(s) \vec{c} and the linear, logarithmic or any transform of the expression ratio \vec{r} across all samples for each individual gene. Specifically, the correlation coefficient may be calculated as:

$$[0090] \quad \rho = (\vec{c} \bullet \vec{r}) / (\|\vec{c}\| \cdot \|\vec{r}\|) \quad \text{Equation (1)}$$

[0091] Markers for which the coefficient of correlation exceeds a cutoff are identified as prognosis-informative markers specific for a particular clinical type, *e.g.*, good prognosis, within a given patient subset. Such a cutoff or threshold may correspond to a certain significance of discriminating genes obtained by Monte Carlo simulations. The threshold depends upon the number of samples used; the threshold can be calculated as $3 \times 1/\sqrt{n-3}$, where $1/\sqrt{n-3}$ is the distribution width and n = the number of samples. In a specific

embodiment, markers are chosen if the correlation coefficient is greater than about 0.3 or less than about -0.3.

[0092] Next, the significance of the correlation is calculated. This significance may be calculated by any statistical means by which such significance is calculated. In a specific example, a set of correlation data is generated using a Monte-Carlo technique to randomize the association between the expression difference of a particular marker and the clinical category. The frequency distribution of markers satisfying the criteria in the Monte-Carlo runs is used to determine whether the number of markers selected by correlation with clinical data is significant.

[0093] Once a marker set is identified, the markers may be rank-ordered in order of significance of discrimination. One means of rank ordering is by the amplitude of correlation between the change in gene expression of the marker and the specific condition being discriminated. Another, preferred, means is to use a statistical metric. In a specific embodiment, the metric is a t-test-like statistic:

$$[0094] \quad t = \frac{(\langle x_1 \rangle - \langle x_2 \rangle)}{\sqrt{[\sigma_1^2(n_1 - 1) + \sigma_2^2(n_2 - 1)] / (n_1 + n_2 - 1) / (1/n_1 + 1/n_2)}} \quad \text{Equation (2)}$$

[0095] In this equation, $\langle x_1 \rangle$ is the error-weighted average of the log ratio of transcript expression measurements within a first clinical group (*e.g.*, good prognosis), $\langle x_2 \rangle$ is the error-weighted average of log ratio within a second, related clinical group (*e.g.*, poor prognosis), σ_1 is the variance of the log ratio within the first clinical group (*e.g.*, good prognosis), n_1 is the number of samples for which valid measurements of log ratios are available, σ_2 is the variance of log ratio within the second clinical group (*e.g.*, poor prognosis), and n_2 is the number of samples for which valid measurements of log ratios are available. The *t*-value represents the variance-compensated difference between two means.

[0096] The rank-ordered marker set may be used to optimize the number of markers in the set used for discrimination. This is accomplished generally in a “leave one out” method as follows. In a first run, a subset, for example five, of the markers from the top of the ranked list is used to generate a template, where out of X samples, X-1 are used to generate the template, and the status of the remaining sample is predicted. This process is repeated for every sample until every one of the X samples is predicted once. In a second run, additional markers, for example five additional markers, are added, so that a template is now generated from 10 markers, and the outcome of the remaining sample is predicted. This process is

repeated until the entire set of markers is used to generate the template. For each of the runs, type 1 error (false negative) and type 2 errors (false positive) are counted; the optimal number of markers is that number where the type 1 error rate, or type 2 error rate, or preferably the total of type 1 and type 2 error rate is lowest.

[0097] For prognostic markers, validation of the marker set may be accomplished by an additional statistic, a survival model. This statistic generates the probability of tumor distant metastases as a function of time since initial diagnosis. A number of models may be used, including Weibull, normal, log-normal, log logistic, log-exponential, or log-Rayleigh (Chapter 12 “Life Testing”, S-PLUS 2000 GUIDE TO STATISTICS, Vol. 2, p. 368 (2000)). For the “normal” model, the probability of distant metastases P at time t is calculated as

$$[0098] \quad P = \alpha \times \exp(-t^2/\tau^2) \quad \text{Equation (3)}$$

[0099] where α is fixed and equal to 1, and τ is a parameter to be fitted and measures the “expected lifetime”.

[00100] It is preferable that the above marker identification process be iterated one or more times by excluding one or more samples from the marker selection or ranking (*i.e.*, from the calculation of correlation). Those samples being excluded are the ones that can not be predicted correctly from the previous iteration. Preferably, those samples excluded from marker selection in this iteration process are included in the classifier performance evaluation, to avoid overstating the performance.

[00101] It will be apparent to those skilled in the art that the above methods, in particular the statistical methods described above, are not limited to the identification of markers associated with the prognosis of breast cancer within a particular patient subset, but may be used to identify set of marker genes associated with any phenotype or condition, or with any subset of a phenotype or condition defined by one or more characteristics of the phenotype or condition. The phenotype or condition can be the presence or absence of a disease such as cancer, or the presence or absence of any identifying clinical condition associated with that cancer. In the disease context, the phenotype may be a prognosis such as a survival time, probability of distant metastases of a disease condition, or likelihood of a particular response to a therapeutic or prophylactic regimen. The phenotype need not be cancer, or a disease; the phenotype may be a nominal characteristic associated with a healthy individual.

[00102] Thus, the invention provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics,

comprising: (a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics of said condition into a plurality of first classes; (b) identifying within each of said first classes a first set of genes or markers informative for said condition, wherein said first set of genes or markers within each of said first classes is unique to said class relative to other classes. In a specific embodiment, samples or individuals in at least one of said first classes are additionally classified on the basis of a phenotypic or genotypic characteristic different from that used to distinguish said first classes into a plurality of second classes, and identifying within at least one of said second classes a second set of informative genes or markers, wherein said second set of informative genes or markers within each of said second classes is unique to said second class relative to other classes. In another embodiment, the invention provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics into a plurality of first classes; (b) classifying at least one of said first classes into a plurality of second classes on the basis of phenotypic or genotypic characteristic different than that used to distinguish said plurality of first classes; (c) identifying within at least one of said first classes or said second classes a set of genes or markers informative for said condition, wherein said set of genes or markers is unique to said class relative to other classes. The invention further provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) selecting a first characteristic from said plurality of phenotypic or genotypic characteristics; (b) identifying at least two first condition classes differentiable by said first characteristic; (c) selecting a plurality of individuals classifiable into at least one of said first condition classes; and (d) identifying in samples derived from each of said plurality of individuals a set of genes or markers informative for said condition within said at least one of said first condition classes.

5.3.3 CLASSIFIER GENESETS FOR FIVE PATIENT SUBSETS

[00103] The present invention provides sets of markers useful for the prognosis of breast cancer. The markers were identified according to the above methods in specific subsets of individuals with breast cancer. Generally, the marker sets were identified within a population of breast cancer patients that had been first stratified into five phenotypic categories based on criteria relevant to breast cancer prognosis, including estrogen receptor (ER) status, lymph

node status, type of mutation(s) (*i.e.*, *BRCA1*-type or sporadic) and age at diagnosis. More specifically, patients, and tumors from which samples were taken, were classified as ER⁻, sporadic (*i.e.*, being both estrogen receptor negative and having a non-*BRCA1*-type tumor); ER⁻, *BRCA1* (*i.e.*, being both estrogen receptor negative and having a *BRCA1*-type tumor); ER⁺, ER/AGE high (*i.e.*, estrogen receptor positive with a high ratio of the log (ratio) of estrogen receptor gene expression to age); ER⁺, ER/AGE low, LN⁺ (*i.e.*, estrogen receptor positive with a low ratio of the log (ratio) of estrogen receptor gene expression to age, lymph node positive); and ER⁺, ER/AGE low, LN⁻ (*i.e.*, estrogen receptor positive with a low ratio of the log (ratio) of estrogen receptor gene expression to age, lymph node negative). The rationale for subdivision of the original patient set into these five subsets is detailed in the Examples (Section 6). The marker sets useful for each of the subsets above are provided in Tables 1-5, respectively.

Table 1: Geneset of 20 markers used to classify ER⁻, sporadic individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list	SEQ ID
AF055033	IGFBP5	-2.12	0.88	0.54	insulin-like growth factor binding protein 5	Growth factor binding, Glycoprotein, Signal, 3D-structure	11
NM_000599	IGFBP5	-3.41	0.43	0.53	insulin-like growth factor binding protein 5	Growth factor binding, Glycoprotein, Signal, 3D-structure	51
L27560	IGFBP5	-4.55	0	0.52	EST	Hypothetical protein	29
AF052162	FLJ12443	-0.27	1.6	0.52	EST	Hypothetical protein	9
NM_001456	FLNA	-0.61	2.47	0.52	filamin A, alpha (actin binding protein 280)	Hypothetical protein, Actin-binding, Phosphorylation, Repeat, Polymorphism, Disease mutation	73
NM_002205	ITGA5	-0.37	2.08	0.49	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	Integrin, Cell adhesion, Receptor, Glycoprotein, Transmembrane, Signal, Calcium, Repeat	93
NM_013261	PPARGC1	0.09	1.54	0.47	peroxisome proliferative activated receptor, gamma, coactivator 1		231
NM_001605	AARS	0.39	2.36	0.51	alanyl-tRNA synthetase	Aminoacyl-tRNA synthetase, Protein biosynthesis, Ligase, ATP-binding	77
X87949	HSPA5	-0.03	2.03	0.49	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	ATP-binding, Hypothetical protein, Endoplasmic reticulum, Signal	273
Contig50950_RC	NGEF	-1.17	3.2	0.52	neuronal guanine nucleotide exchange factor		337

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list	SEQ ID
NM_005689	ABCB6	-0.51	2.26	0.48	ATP-binding cassette, sub-family B (MDR/TAP), member 6	ATP-binding, Transport, Transmembrane, Mitochondrion, Inner membrane, Transit peptide, Hypothetical protein	187
NM_004577	PSPH	-0.56	3.05	0.51	phosphoserine phosphatase	Hydrolase, Serine biosynthesis, Magnesium, Phosphorylation	151
NM_003832	PSPHL	-2.08	2.18	0.5	phosphoserine phosphatase-like		131
NM_002422	MMP3	-0.96	2.54	0.5	matrix metalloproteinase 3 (stromelysin 1, procollagenase)	Hydrolase, Metalloprotease, Glycoprotein, Zinc, Zymogen, Calcium, Collagen degradation, Extracellular matrix, Signal, Polymorphism, 3D-structure	101
Contig37562_RC		-3.42	-6.02	-0.59	ESTs		293
NM_018465	MDS030	-0.82	-3.28	-0.58	uncharacterized hematopoietic stem/progenitor cells protein MDS030	Hypothetical protein	267
Contig54661_RC		-0.79	-2.08	-0.54	ESTs		349
AB032969	KIAA1143	-0.6	-2.85	-0.53	KIAA1143 protein	Hypothetical protein	1
Contig55353_RC	KIAA1915	-0.27	-1.82	-0.47	KIAA1915 protein	Hypothetical protein	353
NM_005213	CSTA	2.11	-3.4	-0.49	cystatin A (stefin A)	Thiol protease inhibitor, 3D-structure	175

Table 2. Geneset of 10 markers used to classify ER⁻, *BRCA1* individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
AF005487		6.08	0.5	-0.79	HLA-DRB6	Homo sapiens MHC class II antigen (DRB6) mRNA, HLA-DRB6*0201 allele, sequence.	MHC	3
Contig50728_RC		4.02	0.25	-0.77		ESTs, Weakly similar to S26650 DNA-binding protein 5 - human [H.sapiens]		333
Contig53598_RC		8.41	3.26	-0.77	FLJ11413	hypothetical protein FLJ11413	Hypothetical protein	343
NM_002888	RARR ES1	6.9	0.05	-0.87	RARRES1	retinoic acid receptor responder (tazarotene induced) 1	Receptor, Transmembrane, Signal-anchor	109

NM_005218	DEFB1	5.14	-3.02	-0.81	DEFB1	defensin, beta 1	Antibiotic, Signal, 3D- structure	177
U17077	BENE	2.72	-1.72	-0.77	BENE	BENE protein	Transmembrane	271
Contig14683_RC		1.29	-2.31	-0.74		ESTs		279
Contig53641_RC		-3.29	4.23	0.75	MAGE-E1	MAGE-E1 protein	Hypothetical protein	345
Contig56678_RC		-6.7	-9.73	-0.82		ESTs, Highly similar to THYA_HUMAN Prothymosin alpha [H.sapiens]		357
NM_005461	KRML	0.88	-3.38	-0.75	MAFB	v-maf musculoaponeuroti c fibrosarcoma oncogene homolog B (avian)	Transcription regulation, Repressor, DNA-binding, Nuclear protein, Hypothetical protein	181

Table 3. Geneset of 50 markers used to classify ER+, ER/AGE high individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
NM_003600	STK15	-2.93	2.08	0.8	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine- protein kinase, Transferase	125
NM_003158	STK6	-1.57	1.42	0.78	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine- protein kinase, Transferase	113
NM_007019	UBCH10	-2.98	2.62	0.81	ubiquitin-conjugating enzyme E2C	Hypothetical protein, Ubl conjugation pathway, Ligase, Multigene family, Mitosis, Cell cycle, Cell division	217
NM_013277	ID-GAP	-2.43	2.43	0.77	Rac GTPase activating protein 1	Hypothetical protein	233
NM_004336	BUB1	-2.04	1.39	0.77	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)	Transferase, Serine/threonine- protein kinase, ATP-binding, Cell cycle, Nuclear protein, Mitosis, Phosphorylation, Polymorphism	147

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
NM_006607	PTTG2	-1.71	1.49	0.72	pituitary tumor-transforming 2		211
AK001166	FLJ11252	-1.33	0.99	0.71	hypothetical protein FLJ11252	Hypothetical protein	13
NM_004701	CCNB2	-4.62	2.01	0.81	cyclin B2	Cyclin, Cell cycle, Cell division, Mitosis	153
Contig57584_RC		-3.68	2.04	0.78	likely ortholog of mouse gene rich cluster, C8 gene		359
NM_006845	KNSL6	-4.13	1.05	0.73	kinesin-like 6 (mitotic centromere-associated kinesin)	Hypothetical protein, Motor protein, Microtubules, ATP-binding, Coiled coil, Nuclear protein	215
Contig38901_RC		-3.08	1.15	0.75	hypothetical protein MGC45866	Hypothetical protein	299
NM_018410	DKFZp762E1312	-4.38	1.49	0.75	hypothetical protein DKFZp762E1312	Hypothetical protein	263
NM_003981	PRC1	-3.52	2.17	0.78	protein regulator of cytokinesis 1		133
NM_001809	CENPA	-5.04	0.98	0.75	centromere protein A, 17kDa	Hypothetical protein, Chromosomal protein, Nuclear protein, DNA-binding, Centromere, Antigen	81
NM_003504	CDC45L	-2.67	1.22	0.73	CDC45 cell division cycle 45-like (S. cerevisiae)	DNA replication, Cell cycle, Nuclear protein, Cell division	123
Contig41413_RC		-5.43	2.15	0.74	ribonucleotide reductase M2 polypeptide	Oxidoreductase, DNA replication, Iron	305
NM_004217	STK12	-2.17	0.73	0.72	serine/threonine kinase 12	Hypothetical protein, ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase	143
NM_002358	MAD2L1	-2.65	2.27	0.83	MAD2 mitotic arrest deficient-like 1 (yeast)	Cell cycle, Mitosis, Nuclear protein, 3D-structure	99
NM_014321	ORC6L	-2.73	1.8	0.75	origin recognition complex, subunit 6 homolog-like (yeast)	Hypothetical protein, DNA replication, Nuclear protein, DNA-binding	241
NM_012291	KIAA0165	-1.52	1.55	0.71	extra spindle poles like 1 (S. cerevisiae)	Hypothetical protein	229

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
NM_004203	PKMYT1	-3.64	2.2	0.7	retinoblastoma-like 2 (p130)	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase, Transcription regulation, DNA-binding, Nuclear protein, Cell cycle, Phosphorylation, Anti-oncogene	137
M96577	E2F1	-2.14	1.42	0.75	E2F transcription factor 1	Transcription regulation, Activator, DNA-binding, Nuclear protein, Phosphorylation, Cell cycle, Apoptosis, Polymorphism	33
NM_002266	KPNA2	-3.77	1.78	0.71	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Transport, Protein transport, Repeat, Nuclear protein, Polymorphism	95
Contig31288_RC		-2.63	0.7	0.68	ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens]		289
NM_014501	E2-EPF	-1.55	1.93	0.7	ubiquitin carrier protein	Ubl conjugation pathway, Ligase, Multigene family	247
NM_001168	BIRC5	-5.76	2.01	0.78	baculoviral IAP repeat-containing 5 (survivin)	Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure, Hypothetical protein, Protease, Receptor	63
NM_003258	TK1	-4.57	1.38	0.71	thymidine kinase 1, soluble	Transferase, Kinase, DNA synthesis, ATP-binding	115
NM_001254	CDC6	-2.46	0.28	0.72	CDC6 cell division cycle 6 homolog (S. cerevisiae)	ATP-binding, Cell division	67
NM_004900	DJ742C19.2	-2.96	0.13	0.69	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	Hydrolase	161
NM_004702	CCNE2	-3.12	2.13	0.81	cyclin E2	Cell cycle, Cell division, Cyclin, Hypothetical protein, Phosphorylation, Alternative splicing, Nuclear protein	155

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
AL160131		-3.07	2.42	0.7	hypothetical protein MGC861	Hypothetical protein	21
NM_016359	LOC5120 3	-3.22	2.61	0.76	nucleolar protein ANKT	Hypothetical protein, Nuclear protein	253
NM_004856	KNSL5	-1.52	1.1	0.71	kinesin-like 5 (mitotic kinesin-like protein 1)	Motor protein, Cell division, Microtubules, ATP- binding, Coiled coil, Mitosis, Cell cycle, Nuclear protein	159
NM_000057	BLM	-1.54	0.76	0.71	Bloom syndrome	Hydrolase, Helicase, ATP- binding, DNA- binding, Nuclear protein, DNA replication, Disease mutation	35
NM_018455	BM039	-2.44	1.18	0.7	uncharacterized bone marrow protein BM039		265
NM_002106	H2AFZ	-2.49	1.53	0.72	H2A histone family, member Z	Chromosomal protein, Nucleosome core, Nuclear protein, DNA-binding, Multigene family	91
Contig64688		-2.68	3.1	0.73	hypothetical protein FLJ23468	Hypothetical protein	365
Contig44289_RC		-1.65	1.6	0.67	ESTs		315
Contig28552_RC		-1.37	1.53	0.68	diaphanous homolog 3 (Drosophila)	Hypothetical protein, Coiled coil, Repeat, Alternative splicing	281
Contig46218_RC		-1.31	1.56	0.68	ESTs, Weakly similar to T19201 hypothetical protein C11G6.3 - Caenorhabditis elegans [C. elegans]		321
Contig28947_RC		-1.3	0.98	0.67	cell division cycle 25A	Hypothetical protein, Cell division, Mitosis, Hydrolase, Alternative splicing, Multigene family, 3D-structure	283
NM_016095	LOC5165 9	-1.4	2.13	0.67	HSPC037 protein	Hypothetical protein	249
NM_003090	SNRPA1	-3.26	0.95	0.7	small nuclear ribonucleoprotein polypeptide A'	Hypothetical protein, Nuclear protein, RNA- binding, Ribonucleoprotein, Leucine-rich repeat, Repeat, 3D-structure	111

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
NM_002811	PSMD7	-2.48	1.89	0.7	proteasome (prosome, macropain) 26S subunit, non-ATPase, 7 (Mov34 homolog)	Proteasome	107
Contig38288_RC		-2.34	0.97	0.67	hypothetical protein DKFZp762A2013	Hypothetical protein	297
NM_003406	YWHAZ	-1.5	2.79	0.68	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	Brain, Neurone, Phosphorylation, Acetylation, Multigene family, 3D-structure	121
AL137540	NTN4	2.13	-4.61	-0.69	netrin 4	Hypothetical protein, Laminin EGF-like domain, Signal	19
AL049367		1.9	-3.2	-0.68	EST	Transducer, Prenylation, Lipoprotein, Multigene family, Acetylation	15
NM_013409	FST	1.04	-5.78	-0.69	folistatin	Glycoprotein, Repeat, Signal, Alternative splicing	235
NM_000060	BTD	3.1	-1.45	-0.67	biotinidase	Hydrolase, Glycoprotein, Signal, Disease mutation	37

Table 4. Geneset of 50 markers used to classify ER+, ER/AGE low, LN+ individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
--------------------------	------	---------------------	---------------------	------------------	-------------	--------------------------	--------

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
NM_006417	MTAP44	-1.5	3	0.69	Fc fragment of IgG, low affinity IIb, receptor for (CD32)	Hydrolase, Hypothetical protein, Immunoglobulin domain, IgG- binding protein, Receptor, Transmembrane, Glycoprotein, Signal, Repeat, Multigene family, Polymorphism, NAD, One-carbon metabolism, Serine protease, Zymogen, Protease, Alternative splicing, Chromosomal translocation, Proto-oncogene, Galactin, Lectin, Antigen	205
NM_006820	GS3686	-4.3	4.06	0.69	chromosome 1 open reading frame 29	Hypothetical protein	213
NM_001548	IFIT1	-3.4	4.27	0.71	Interferon-induced protein with tetra-tryptophan repeats 1	Repeat, TPR repeat, Interferon induction	75
Contig41538_RC		-2.5	3.16	0.68	ESTs, Moderately similar to hypothetical protein FLJ20489 [<i>Homo sapiens</i>]		307
NM_016816	OAS1	-1.7	3.29	0.75	2',5'-oligoadenylate synthetase 1, 40/46kDa	RNA-binding, Transferase, Nucleotidyltransfer ase, Interferon induction, Alternative splicing	255
Contig51660_RC		-2.1	2.65	0.66	28kD interferon responsive protein	Transmembrane	339
Contig43645_RC		-4.8	1.44	0.63	<i>Homo sapiens</i> , clone IMAGE:4428577, mRNA, partial cds	Hypothetical protein	313
AF026941		-4.6	2.71	0.63	EST, Weakly similar to 2004399A chromosomal protein [<i>Homo sapiens</i>]	Hypothetical protein	5
NM_007315	STAT1	-3.5	1.8	0.59	signal transducer and activator of transcription 1, 91kDa	Transcription regulation, DNA- binding, Nuclear protein, Phosphorylation, SH2 domain, Alternative splicing, 3D-structure	225

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
NM_002038	G1P3	-4.1	5.64	0.79	interferon, alpha-inducible protein (clone IFI-6-16)	Interferon induction, Transmembrane, Signal, Alternative splicing	85
NM_005101	ISG15	-5.6	5.34	0.77	interferon-stimulated protein, 15 kDa	Interferon induction, Repeat	169
NM_002462	MX1	-6.1	0.83	0.56	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	Hypothetical protein, Interferon induction, GTP-binding, Multigene family, Antiviral	103
NM_005532	IFI27	-5.8	2.81	0.59	interferon, alpha-inducible protein 27	Interferon induction, Transmembrane	183
NM_002346	LY6E	-2.1	3.58	0.75	lymphocyte antigen 6 complex, locus E	Signal, Antigen, Multigene family, Membrane, GPI-anchor	97
NM_016817	OAS2	-3.6	1.89	0.59	2'-5'-oligoadenylate synthetase 2, 69/71kDa	RNA-binding, Transferase, Nucleotidyltransferase, Repeat, Interferon induction, Alternative splicing, Myristate	257
Contig44909_RC		-2.3	1.13	0.55	hypothetical protein BC012330	Hypothetical protein	317
NM_017414	USP18	-4.1	3.37	0.72	ubiquitin specific protease 18	Ubl conjugation pathway, Hydrolase, Thiol protease, Multigene family	259
NM_004029	IRF7	-2.4	3.67	0.66	interferon regulatory factor 7	Collagen, Transcription regulation, DNA-binding, Nuclear protein, Activator, Alternative splicing	135
NM_004335	BST2	-3.2	3.22	0.57	bone marrow stromal cell antigen 2	Transmembrane, Glycoprotein, Signal-anchor, Polymorphism	145
NM_002759	PRKR	-2.4	1.8	0.58	protein kinase, interferon-inducible double stranded RNA dependent	Transferase, Serine/threonine-protein kinase, ATP-binding, Repeat, Phosphorylation, Interferon induction, RNA-binding, 3D-structure	105

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
NM_006332	IFI30	-3.8	2.65	0.64	interferon, gamma-inducible protein 30	Oxidoreductase, Interferon induction, Glycoprotein, Lysosome, Signal, Hypothetical protein	203
NM_009587	LGALS9	-3.2	2.08	0.6	lectin, galactoside-binding, soluble, 9 (galectin 9)	Galaptin, Lectin, Repeat, Alternative splicing	227
NM_003641	IFITM1	-2.4	5.54	0.63	interferon induced transmembrane protein 1 (9-27)	Interferon induction, Transmembrane	127
NM_017523	HSXIAPAF1	-1	2.84	0.7	XIAP associated factor-1	Hypothetical protein	261
NM_014314	RIG-I	-1.3	3.55	0.62	RNA helicase	ATP-binding, Helicase, Hydrolase, Hypothetical protein	239
Contig47563_RC		-2.2	3.11	0.56	ESTs		325
AI497657_RC		-4.4	5.61	0.74	guanine nucleotide binding protein 4	Transducer, Prenylation, Lipoprotein, Multigene family	335
NM_000735	CGA	-4.3	2.5	0.58	glycoprotein hormones, alpha polypeptide	Hormone, Glycoprotein, Signal, 3D-structure	53
NM_004988	MAGEA1	-1.4	6.31	0.64	melanoma antigen, family A, 1 (directs expression of antigen MZ2-E)	Antigen, Multigene family, Polymorphism, Tumor antigen	163
Contig54242_RC		-1.2	4.1	0.65	chromosome 17 open reading frame 26	Hypothetical protein	347
NM_004710	SYNGR2	-1.4	3.01	0.54	synaptogyrin 2	Transmembrane	157
NM_001168	BIRC5	-3.7	3.39	0.64	baculoviral IAP repeat-containing 5 (survivin)	Hypothetical protein, Protease, Receptor, Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure	63
Contig41413_RC		-4.4	2.61	0.57	ribonucleotide reductase M2 polypeptide	Oxidoreductase, DNA replication, Iron	305

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
NM_004203	PKMYT1	-3.4	3.79	0.6	retinoblastoma-like 2 (p130)	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase, Transcription regulation, DNA-binding, Nuclear protein, Cell cycle, Phosphorylation, Anti-oncogene	137
Contig48913_RC		-3.1	1.72	0.55	<i>Homo sapiens</i> , Similar to hypothetical protein PRO1722, clone MGC:15692 IMAGE:3351479, mRNA, complete cds		327
NM_005804	DDXL	-2.5	1.42	0.58	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 39	ATP-binding, Helicase, Hydrolase, Hypothetical protein	191
NM_016359	LOC51203	-1.7	3.6	0.57	nucleolar protein ANKT	Hypothetical protein, Nuclear protein	253
NM_001645	APOC1	-2.9	3.43	0.58	apolipoprotein C-I	Plasma, Lipid transport, VLDL, Signal, 3D-structure, Polymorphism	79
Contig37895_RC		-2	2.05	0.55	ESTs		295
NM_005749	TOB1	-1.3	4.96	0.59	transducer of ERBB2, 1	Phosphorylation	189
NM_000269	NME1	-1.3	2.98	0.55	non-metastatic cells 1, protein (NM23A) expressed in	Transferase, Kinase, ATP-binding, Nuclear protein, Anti-oncogene, Disease mutation	39
NM_014462	LSM1	-1	4.5	0.57	Lsm1 protein	Nuclear protein, Ribonucleoprotein, mRNA splicing, mRNA processing, RNA-binding	245
Contig31221_RC		-1.4	3.83	0.56	HTPAP protein		287
NM_005326	HAGH	-1.9	4.29	0.57	hydroxyacyl glutathione hydrolase	Hydrolase, Zinc, 3D-structure	179
Contig42342_RC		0.78	-3.2	-0.6	<i>Homo sapiens</i> cDNA FLJ39417 fis, clone PLACE6016942	Hypothetical protein	311
AL137540	NTN4	2.24	-3.9	-0.6	netrin 4	Laminin EGF-like domain, Signal, Hypothetical protein	19
Contig40434_RC		1.64	-5.6	-0.6	wingless-type MMTV integration site family, member 5A	Developmental protein, Glycoprotein, Signal	301

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
Contig1632_RC		1.03	-3.9	-0.6	hypothetical protein MGC17921	Hypothetical protein	275
NM_014246	CELSR1	0.95	-4.6	-0.6	cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, <i>Drosophila</i>)	G-protein coupled receptor, Transmembrane, Glycoprotein, EGF- like domain, Calcium-binding, Laminin EGF-like domain, Repeat, Developmental protein, Hydroxylation, Signal, Alternative splicing, Hypothetical protein	237
NM_005139	ANXA3	1.26	-6.2	-0.6	annexin A3	Annexin, Calcium/phospholi pid-binding, Repeat, Phospholipase A2 inhibitor, 3D- structure, Polymorphism	171

Table 5. Geneset of 65 markers used to classify ER+, ER/AGE low, LN⁻ individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
M55914	MPB1	-2.82	1.25	0.5	ENO1	enolase 1, (alpha)	DNA-binding, Transcription regulation, Repressor, Nuclear protein, Lyase, Glycolysis, Magnesium, Multigene family, Hypothetical protein	31
NM_005945	MPB1	-3.06	1.19	0.49	ENO1	Homo sapiens enolase 1, (alpha) (ENO1), mRNA.	Glycolysis, Hypothetical protein, Lyase, Magnesium, DNA-binding, Transcription regulation, Repressor, Nuclear protein, Multigene family	193

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_001428	ENO1	-2.53	1.18	0.46	ENO1	enolase 1, (alpha)	DNA-binding, Transcription regulation, Repressor, Nuclear protein, Lyase, Glycolysis, Magnesium, Multigene family, Hypothetical protein	71
NM_001216	CA9	-4.72	1.49	0.6	CA9	carbonic anhydrase IX	Lyase, Zinc, Transmembrane , Glycoprotein, Antigen, Signal, Nuclear protein, Polymorphism	65
NM_001124	ADM	-5.68	2.99	0.56	ADM	Adrenomedullin	Hormone, Amidation, Cleavage on pair of basic residues, Signal	61
NM_000584	IL8	-2.45	2.04	0.54	IL8	interleukin 8	Cytokine, Chemotaxis, Inflammatory response, Signal, Alternative splicing, 3D- structure	49
D25328	PFKP	-4.19	3.29	0.56	PFKP	Phosphofructo- kinase, platelet	Kinase, Transferase, Glycolysis, Repeat, Allosteric enzyme, Phosphorylation, Magnesium, Multigene family	25
NM_006096	NDRG1	-5.45	5.97	0.77	NDRG1	N-myc downstream regulated gene 1	Hypothetical protein, Nuclear protein, Repeat	199
NM_004994	MMP9	-5.53	1.07	0.49	MMP9	matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	Hydrolase, Metalloprotease, Glycoprotein, Zinc, Zymogen, Calcium, Collagen degradation, Extracellular matrix, Repeat, Signal, Polymorphism, 3D-structure	165

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_003311	TSSC3	-4.57	5.58	0.68	TSSC3	tumor suppressing subtransferable candidate 3		117
NM_006086	TUBB4	-5.19	2.85	0.59	TUBB4	tubulin, beta, 4	G-protein coupled receptor, Transmembrane, Glycoprotein, Phosphorylation, Lipoprotein, Palmitate, Polymorphism, Hypothetical protein, GTP-binding, Receptor, Microtubules, Multigene family	197
NM_006115	PRAME	-4.48	2.77	0.61	PRAME	preferentially expressed antigen in melanoma	Antigen	201
NM_004345	CAMP	-2.02	1.37	0.49	CAMP	cathelicidin antimicrobial peptide	Antibiotic, Signal	149
NM_018455	BM039	-2.34	0.76	0.47	BM039	uncharacterized bone marrow protein BM039		265
Contig49169_RC		-1.17	1.5	0.46	SUV39H2	suppressor of variegation 3-9 (Drosophila) homolog 2; hypothetical protein FLJ23414	Hypothetical protein, Nuclear protein	329
Contig45032_RC		-1.37	0.77	0.45	FLJ14813	hypothetical protein FLJ14813	Hypothetical protein, ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase	319
NM_000917	P4HA1	-1.54	4.31	0.62	P4HA1	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide I	Dioxygenase, Collagen, Oxidoreductase, Iron, Vitamin C, Alternative splicing, Glycoprotein, Endoplasmic reticulum, Signal	57
NM_002046	GAPD	-2.51	3.42	0.6	GAPD	glyceraldehyde-3-phosphate dehydrogenase	Glycolysis, NAD, Oxidoreductase, Hypothetical protein, Multigene family	87

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_000365	TPI1	-1.81	2.94	0.56	TPI1	triosephosphate isomerase 1	Fatty acid biosynthesis, Gluconeogenesi s, Glycolysis, Isomerase, Pentose shunt, Disease mutation, Polymorphism, 3D-structure	45
NM_014364	GAPDS	-1.08	2.88	0.58	GAPDS	glyceraldehyde-3- phosphate dehydrogenase, testis-specific	Glycolysis, Oxidoreductase, NAD	243
NM_005566	LDHA	-2.01	4.01	0.59	LDHA	lactate dehydrogenase A	Oxidoreductase, NAD, Glycolysis, Multigene family, Disease mutation, Polymorphism	185
NM_000291	PGK1	-2.28	1.68	0.51	PGK1	phosphoglycerate kinase 1	Kinase, Transferase, Multigene family, Glycolysis, Acetylation, Disease mutation, Polymorphism, Hereditary hemolytic anemia	41
NM_016185	LOC511 55	-2.33	2.82	0.59	HN1	hematological and neurological expressed 1		251
NM_001168	BIRC5	-4.33	2.78	0.55	BIRC5	baculoviral IAP repeat-containing 5 (survivin)	Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D- structure, Hypothetical protein, Protease, Receptor	63
NM_002266	KPNA2	-3.75	1.34	0.47	KPNA2	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Transport, Protein transport, Repeat, Nuclear protein, Polymorphism	95
Contig31288_RC		-2.1	1.27	0.5		ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens]		289

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_000269	NME1	-2.15	3.43	0.55	NME1	non-metastatic cells 1, protein (NM23A) expressed in	Transferase, Kinase, ATP- binding, Nuclear protein, Anti- oncogene, Disease mutation	39
NM_003158	STK6	-1.23	1.73	0.45	STK6	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine -protein kinase, Transferase	113
NM_007274	HBACH	-1.83	2.73	0.51	BACH	brain acyl-CoA hydrolase	Hydrolase, Serine esterase, Repeat	223
Contig55188_RC		-2.36	3.28	0.47	FLJ22341	hypothetical protein FLJ22341	Hypothetical protein	351
NM_002061	GCLM	-1.06	1.76	0.48	GCLM	glutamate-cysteine ligase, modifier subunit	Ligase, Glutathione biosynthesis	89
NM_004207	SLC16A 3	-3.11	5.07	0.67	SLC16A3	solute carrier family 16 (monocarboxylic acid transporters), member 3	Transport, Symport, Transmembrane , Multigene family	139
NM_000582	SPP1	-5.09	5.47	0.53	SPP1	secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	Hypothetical protein, Glycoprotein, Sialic acid, Biom mineralizatio n, Cell adhesion, Phosphorylation, Signal, Alternative splicing	47
NM_001109	ADAM8	-2.5	3.74	0.45	ADAM8	a disintegrin and metalloproteinase domain 8	Hydrolase, Metalloprotease, Zinc, Signal, Glycoprotein, Transmembrane , Antigen	59
D50402	SLC11A 1	-1.05	3.46	0.53	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1	Transport, Iron transport, Transmembrane , Glycoprotein, Macrophage, Polymorphism	27
AL080235	DKFZP5 86E162 1	-1.23	1.96	0.51	RIS1	Ras-induced senescence 1	Hypothetical protein	17
Contig40552_RC		-1.26	3.96	0.54	FLJ25348	hypothetical protein FLJ25348	Hypothetical protein	303
Contig52490_RC		-0.64	3.33	0.61	LOC11623 8	hypothetical protein BC014072		341
NM_006461	DEEPE ST	-2.1	1.85	0.46	SPAG5	sperm associated antigen 5	Hypothetical protein	207

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
Contig56503_RC		-4.3	3.39	0.55	MGC9753	hypothetical gene MGC9753	Hypothetical protein	355
Contig63525		-1.91	3.34	0.5	FLJ13352	hypothetical protein FLJ13352	Hypothetical protein	363
NM_001909	CTSD	-0.83	4.6	0.51	CTSD	cathepsin D (lysosomal aspartyl protease)	Hydrolase, Aspartyl protease, Glycoprotein, Lysosome, Signal, Zymogen, Polymorphism, Alzheimer's disease, 3D- structure	83
NM_005063	SCD	-2.57	5.15	0.48	SCD	stearoyl-CoA desaturase (delta- 9-desaturase)	Hypothetical protein, Endoplasmic reticulum, Fatty acid biosynthesis, Iron, Oxidoreductase, Transmembrane	167
NM_005165	ALDOC	-2.43	5.02	0.48	ALDOC	aldolase C, fructose- bisphosphate	Lyase, Schiff base, Glycolysis, Multigene family	173
NM_000363	TNNI3	-0.54	3.58	0.48	TNNI3	troponin I, cardiac	Hypothetical protein, Muscle protein, Actin- binding, Acetylation, Disease mutation, Cardiomyopathy , Receptor, Signal	43
AF035284		-1.63	3.28	0.47	FADS1	EST	Heme, Hypothetical protein	7
Contig30875_RC		-0.88	3	0.6		ESTs		285
NM_018487	HCA112	-0.7	3.54	0.58	HCA112	hepatocellular carcinoma- associated antigen 112	Hypothetical protein	269
NM_001323	CST6	-1.63	3.84	0.57	CST6	cystatin E/M	Thiol protease inhibitor, Signal, Glycoprotein	69

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_006516	SLC2A1	-1.66	2.22	0.46	SLC2A1	solute carrier family 2 (facilitated glucose transporter), member 1	Transmembrane, Sugar transport, Transport, Glycoprotein, Multigene family, Disease mutation	209
NM_007267	LAK-4P	-1.04	3.28	0.61	EVIN1	expressed in activated T/LAK lymphocytes	Hypothetical protein	221
NM_004710	SYNGR2	-0.84	4.81	0.56	SYNGR2	synaptogyrin 2	Transmembrane	157
Contig63649_RC		-1.34	6.3	0.75		ESTs, Weakly similar to 2004399A chromosomal protein [Homo sapiens] [H.sapiens]		361
NM_003376	VEGF	-2.12	2.42	0.46	VEGF	vascular endothelial growth factor	Hypothetical protein, Mitogen, Angiogenesis, Growth factor, Glycoprotein, Signal, Heparin-binding, Alternative splicing, Multigene family, 3D-structure	119
NM_000799	EPO	-0.75	4.01	0.69	EPO	erythropoietin	Erythrocyte maturation, Glycoprotein, Hormone, Signal, Pharmaceutical, 3D-structure	55
NM_006014	DXS9879E	-1.85	3.44	0.54	DXS9879E	DNA segment on chromosome X (unique) 9879 expressed sequence		195
NM_007183	PKP3	-0.91	4.14	0.48	PKP3	plakophilin 3	Cell adhesion, Cytoskeleton, Structural protein, Nuclear protein, Repeat	219
D13642	SF3B3	-0.65	2.28	0.48	SF3B3	splicing factor 3b, subunit 3, 130kDa	Hypothetical protein, Spliceosome, mRNA processing, mRNA splicing, Nuclear protein	23
NM_003756	EIF3S3	-1.85	2.19	0.46	EIF3S3	eukaryotic translation initiation factor 3, subunit 3 gamma, 40kDa	Initiation factor, Protein biosynthesis	129

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
Contig47096_RC		-0.41	4.52	0.54	PFKFB4	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4	Kinase, Multifunctional enzyme, Transferase, Hydrolase, ATP-binding, Phosphorylation, Multigene family	323
NM_004209	SYNGR3	-0.31	3.67	0.53	SYNGR3	synaptogyrin 3	Transmembrane	141
Contig3464_RC		0.99	-5.81	-0.52		ESTs		277
Contig31646_RC		1.1	-7.76	-0.5	COL14A1	collagen, type XIV, alpha 1 (undulin)	Extracellular matrix, Glycoprotein, Hypothetical protein, Collagen, Signal	291
Contig49388_RC		1.73	-1.75	-0.51	FLJ13322	hypothetical protein FLJ13322	Hypothetical protein	331
Contig41887_RC		0.37	-5.74	-0.47	LOC124220	similar to common salivary protein 1	Hypothetical protein	309

[00104]

5.4 DIAGNOSTIC AND PROGNOSTIC METHODS

5.4.1 SAMPLE COLLECTION

[00105] In the present invention, markers, such as target polynucleotide molecules or proteins, are extracted from a sample taken from an individual afflicted with a condition such as breast cancer. The sample may be collected in any clinically acceptable manner, but must be collected such that marker-derived polynucleotides (*i.e.*, RNA) are preserved (if gene expression is to be measured) or proteins are preserved (if encoded proteins are to be measured). For example, mRNA or nucleic acids derived therefrom (*i.e.*, cDNA or amplified DNA) are preferably labeled distinguishably from standard or control polynucleotide molecules, and both are simultaneously or independently hybridized to a microarray comprising some or all of the markers or marker sets or subsets described above.

Alternatively, mRNA or nucleic acids derived therefrom may be labeled with the same label as the standard or control polynucleotide molecules, wherein the intensity of hybridization of each at a particular probe is compared. A sample may comprise any clinically relevant tissue sample, such as a tumor biopsy or fine needle aspirate, or a sample of bodily fluid, such as blood, plasma, serum, lymph, ascitic fluid, cystic fluid, urine or nipple exudate. The sample may be taken from a human, or, in a veterinary context, from non-human animals such as ruminants, horses, swine or sheep, or from domestic companion animals such as felines and canines.

[00106] Methods for preparing total and poly(A)+ RNA are well known and are described generally in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)) and Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994)).

[00107] RNA may be isolated from eukaryotic cells by procedures that involve lysis of the cells and denaturation of the proteins contained therein. Cells of interest include wild-type cells (*i.e.*, non-cancerous), drug-exposed wild-type cells, tumor- or tumor-derived cells, modified cells, normal or tumor cell line cells, and drug-exposed modified cells. Preferably, the cells are breast cancer tumor cells.

[00108] Additional steps may be employed to remove DNA. Cell lysis may be accomplished with a nonionic detergent, followed by microcentrifugation to remove the nuclei and hence the bulk of the cellular DNA. In one embodiment, RNA is extracted from cells of the various types of interest using guanidinium thiocyanate lysis followed by CsCl centrifugation to separate the RNA from DNA (Chirgwin *et al.*, *Biochemistry* 18:5294-5299 (1979)). Poly(A)+ RNA is selected by selection with oligo-dT cellulose (*see* Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)). Alternatively, separation of RNA from DNA can be accomplished by organic extraction, for example, with hot phenol or phenol/chloroform/isoamyl alcohol.

[00109] If desired, RNase inhibitors may be added to the lysis buffer. Likewise, for certain cell types, it may be desirable to add a protein denaturation/digestion step to the protocol.

[00110] For many applications, it is desirable to preferentially enrich mRNA with respect to other cellular RNAs, such as transfer RNA (tRNA) and ribosomal RNA (rRNA). Most mRNAs contain a poly(A) tail at their 3' end. This allows them to be enriched by affinity chromatography, for example, using oligo(dT) or poly(U) coupled to a solid support, such as cellulose or Sephadex™ (*see* Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994)). Once bound, poly(A)+ mRNA is eluted from the affinity column using 2 mM EDTA/0.1% SDS.

[00111] The sample of RNA can comprise a plurality of different mRNA molecules, each different mRNA molecule having a different nucleotide sequence. In a specific embodiment, the mRNA molecules in the RNA sample comprise at least 5, 10, 15, 20, 25, 30, 40 or 50 different nucleotide sequences. More preferably, the mRNA molecules of the RNA sample

comprise mRNA molecules corresponding to each of the marker genes. In another specific embodiment, the RNA sample is a mammalian RNA sample.

[00112] In a specific embodiment, total RNA or mRNA from cells are used in the methods of the invention. The source of the RNA can be cells of a plant or animal, human, mammal, primate, non-human animal, dog, cat, mouse, rat, bird, yeast, eukaryote, prokaryote, etc. In specific embodiments, the method of the invention is used with a sample containing total mRNA or total RNA from 1×10^6 cells or less. In another embodiment, proteins can be isolated from the foregoing sources, by methods known in the art, for use in expression analysis at the protein level.

[00113] Probes to the homologs of the marker sequences disclosed herein can be employed preferably when non-human nucleic acid is being assayed.

[00114] The methods of the invention may employ any molecule suitable as a marker. For example, sets of proteins informative for a particular condition, including a disease, may be determined. As for gene-based markers, levels of variations of different proteins in samples may be determined for phenotypic or genotypic subsets of the condition, and proteins showing significant variation in either level (abundance) or activity, or both, may be identified in order to create a set of proteins informative for one or more of these subsets. Such proteins may be identified, for example, by use of gel electrophoresis, such as one-dimensional polyacrylamide gel electrophoresis, two-dimensional polyacrylamide gel electrophoresis, nondenaturing polyacrylamide gel electrophoresis; isoelectric focusing gels, etc., by use of antibody arrays, etc. Of course, the particular template(s) used to classify the individual depends upon the type(s) of cellular constituents used as markers. For example, where nucleic acids (*e.g.*, genes or nucleic acids derived from expressed genes) are used as markers, the template comprises nucleic acids (or the level of expression or abundance thereof); where proteins are used as markers, the template comprises proteins, for example, the level or abundance of those proteins in a set of individuals; etc.

5.4.2 USE OF PROGNOSTIC GENESETS FOR BREAST CANCER

[00115] According to the present invention, once genesets informative for a plurality of subsets of a condition are identified, an individual is classified into one of these subsets and a prognosis is made based on the expression of the genes, or their encoded proteins, in the geneset for that subset in a breast cancer tumor sample taken from the individual.

[00116] For example, a particular hypothetical condition has three relevant phenotypic characteristics, A, B and C. In this example, based on these characteristics, genesets

informative for prognosis of four patient subsets A^+B^+ ; $A^+B^-C^+$; $A^+B^-C^-$; and A^- are identified by the method described above. Thus, an individual having the condition would first be classified according to phenotypes A-C into one of the four patient subsets. In one embodiment, therefore, the invention provides for the classification of an individual having a condition into one of a plurality of patient subsets, wherein a set of genes informative for prognosis for the subset has been identified. A sample is then taken from the individual, and the expression of the prognostically-informative genes in the sample is analyzed and compared to a control. In various embodiments, the control is the average expression of informative genes in a pool of samples taken from good prognosis individuals classifiable into that patient subset; the average expression of informative genes in a pool of samples taken from poor prognosis individuals classifiable into that patient subset; a set of mathematical values that represent gene expression levels of good prognosis individuals classifiable into that patient subset; etc.

[00117] In another embodiment, a sample is taken from the individual, and the levels of expression of the prognostically-informative genes in the sample is analyzed. In one embodiment, the expression level of each gene can be compared to the expression level of the corresponding gene in a control of reference sample to determine a differential expression level. The expression profile comprising expression levels or differential expression levels of the plurality of genes is then compared to a template profile. In various embodiments, the template profile is a good prognosis template comprising the average expression of informative genes in samples taken from good prognosis individuals classifiable into that patient subset; or a poor prognosis template comprising the average expression of informative genes in samples taken from poor prognosis individuals classifiable into that patient subset; or a good prognosis profile comprising a set of mathematical values that represent gene expression levels of good prognosis individuals classifiable into that patient subset; etc.

[00118] In a specific embodiment, the condition is breast cancer, and the phenotypic, genotypic and/or clinical classes are: ER^- , *BRCA1* individuals; ER^- , sporadic individuals; ER^+ , *ER/AGE* high individuals; ER^+ , *ER/AGE* low, *LN+* individuals; and ER^+ , *ER/AGE* low, *LN^-* individuals. In this embodiment, an individual may be classified as ER^+ or ER^- . If the individual is ER^- , the individual is additionally classified as having a *BRCA1*-type or sporadic tumor. ER^- individuals are thus classified as ER^- , *BRCA1* or ER^- , sporadic. Alternatively, if the individual is classified as ER^+ , the individual is additionally classified as having a high or low ratio of the log (ratio) of the level of expression of the gene encoding the estrogen receptor to the individual's age. Individuals having a low ratio are additionally

classified as LN+ or LN-. ER+ individuals are thus classified as ER+, ER/AGE high; ER+, ER/AGE low, LN+, or ER+, ER/AGE low, LN-. Of course, the individual's ER status, tumor type, age and LN status may be identified in any order, as long as the individual is classified into one of these five subsets.

[00119] Thus, in one embodiment, the invention provides a method of classifying an individual with a condition as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual into one of a plurality of patient classes, said patient classes being differentiated by one or more phenotypic, genotypic or clinical characteristics of said condition; (b) determining the level of expression of a plurality of genes or their encoded proteins in a cell sample taken from the individual relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins in a cell sample taken from the individual relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins informative for prognosis of the patient class into which said individual is classified; and (c) classifying said individual as having a good prognosis or a poor prognosis on the basis of said level of expression. In a specific embodiment, said condition is breast cancer, said good prognosis is the non-occurrence of metastases within five years of initial diagnosis, and said poor prognosis is the occurrence of metastases within five years of initial diagnosis. In an more specific embodiment, said classifying said individual with a condition as having a good prognosis or a poor prognosis is carried out by comparing the level expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a good prognosis or poor prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in a good prognosis control or a poor prognosis control, respectively, more strongly than would be expected by chance. In a more specific embodiment of the method, said plurality of patient subsets comprises ER⁻, *BRCA1* individuals; ER⁻, sporadic individuals; ER+, ER/AGE high individuals; ER+, ER/AGE low, LN+ individuals; and ER+, ER/AGE low, LN⁻ individuals. In another embodiment, said control is the average level of expression of each of said plurality of genes informative for prognosis in a pool of tumor samples from individuals classified into said subset who have a good prognosis or good outcome, or who have a poor prognosis or good outcome. In another specific embodiment, said control is a set of mathematical values representing the average level of expression of genes informative for prognosis in tumor samples of individuals classifiable into said subset who have a good prognosis, or who have a poor prognosis.

[00120] It is evident that the different patient subsets described herein reflect different molecular mechanisms of the initiation of tumor formation and metastasis. Thus, the genesets listed in tables 1-5 are also useful for diagnosing a person as having a particular type of breast cancer in the first instance. Thus, the invention also provides a method of diagnosing an individual as having a particular subtype of breast cancer, comprising determining the level of expression in a sample from said individual of a plurality of the genes for which markers are listed in Tables 1-5; and comparing said expression to a control, where said control is representative of the expression of said plurality of genes in a breast cancer sample of said subtype of cancer, and on the basis of said comparison, diagnosing the individual as having said subtype of breast cancer. In a specific embodiment, said subtype of cancer is selected from the group consisting of ER⁻, *BRCA1* type; ER⁻, sporadic type; ER⁺, ER/AGE high type; ER⁺, ER/AGE low, LN⁺ type; and ER/AGE low, LN⁻ type. In another specific embodiment, said control is the average level of expression of a plurality of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5. In another specific example, said comparing comprises determining the similarity of the expression of the genes for which markers are listed in each of Tables 1-5 in said sample taken from said individual to a control level of expression of the same genes for each of Tables 1-5, and determining whether the level of expression of said genes in said sample is most similar to said control expression of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5.

[00121] In another embodiment, the invention provides a method of classifying an individual as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual as ER⁻, *BRCA1*; ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining the level of expression of a first plurality of genes in a cell sample taken from the individual relative to a control, said first plurality of genes comprising two of the genes corresponding to the markers Table 1 if said individual is classified as ER⁻, *BRCA1*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻, wherein said individual is “ER/AGE high” if the ratio of ER expression to age exceeds a predetermined value, and “ER/AGE low” if the ratio of ER expression to age does not exceed said predetermined value. In a specific embodiment of this method, said predetermined value of ER calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said individual. In another specific embodiment, said individual is ER⁻, *BRCA1*, and said

plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 1. In another specific embodiment, said individual is ER⁻, sporadic, and said plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 2. In another specific embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 3. In another specific embodiment, said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 5. In another specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

[00122] In one embodiment, the invention provides a method of classifying an individual with a condition as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual into one of a plurality of patient classes, said patient classes being differentiated by one or more phenotypic, genotypic or clinical characteristics of said condition; (b) determining the levels of expression of a plurality of genes or their encoded proteins in a cell sample taken from the individual, optionally relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins informative for prognosis of the patient class into which said individual is classified; and (c) classifying said individual as having a good prognosis or a poor prognosis on the basis of said levels of expression. In a specific embodiment, said condition is breast cancer, said good prognosis is the non-occurrence of metastases within five years of initial diagnosis, and said poor prognosis is the occurrence of metastases within five years of initial diagnosis. In a more specific embodiment, said classifying said individual with a condition as having a good prognosis or a poor prognosis is carried out by comparing the patient's expression profile of said plurality of genes or their encoded proteins to a good and/or poor prognosis template profile of expression levels of said plurality of genes or their encoded proteins, and classifying said individual as having a good prognosis or poor prognosis if said patient expression profile has a high similarity to a good prognosis template or a poor prognosis template, respectively. In a more specific embodiment of the method, said plurality of patient subsets comprises ER⁻, *BRCA1* individuals; ER⁻, sporadic individuals; ER⁺, ER/AGE high

individuals; ER+, ER/AGE low, LN+ individuals; and ER+, ER/AGE low, LN⁻ individuals. In another embodiment, said good prognosis template comprises the average level of expression of each of said plurality of genes informative for prognosis in tumor samples from individuals classified into said subset who have a good prognosis or good outcome, while said poor prognosis template comprises the average level of expression of each of said plurality of genes informative for prognosis in tumor samples from individuals classified into said subset who have a poor prognosis or poor outcome. In another specific embodiment, said good or poor prognosis template is a set of mathematical values representing the average level of expression of genes informative for prognosis in tumor samples of individuals classifiable into said subset who have a good prognosis, or who have a poor prognosis, respectively.

[00123] It is evident that the different patient subsets described herein reflect different molecular mechanisms of the initiation of tumor formation and metastasis. Thus, the genesets listed in tables 1-5 are also useful for diagnosing a person as having a particular type of breast cancer in the first instance. Thus, the invention also provides a method of diagnosing an individual as having a particular subtype of breast cancer, comprising determining an expression profile of a plurality of the genes for which markers are listed in Tables 1-5 in a sample from said individual; and comparing said expression profile to a template profile, where said template is representative of the expression of said plurality of genes in a breast cancer sample of said subtype of cancer, and on the basis of said comparison, diagnosing the individual as having said subtype of breast cancer. In a specific embodiment, said subtype of cancer is selected from the group consisting of ER⁻, *BRCA1* type; ER⁻, sporadic type; ER+, ER/AGE high type; ER+, ER/AGE low, LN+ type; and ER/AGE low, LN⁻ type. In another specific embodiment, said template comprises the average levels of expression of a plurality of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5. In another specific example, said comparing comprises determining the similarity of the expression profile of the genes for which markers are listed in each of Tables 1-5 in said sample taken from said individual to a template profile comprising levels of expression of the same genes for each of Tables 1-5, and determining whether the pattern of expression of said genes in said sample is most similar to the pattern of expression of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5.

[00124] In another embodiment, the invention provides a method of classifying an individual as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual as

ER⁻, *BRCA1*; ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining an expression profile of a first plurality of genes in a cell sample taken from the individual relative to a control, said first plurality of genes comprising at least two of the genes corresponding to the markers Table 1 if said individual is classified as ER⁻, *BRCA1*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻, wherein said individual is “ER/AGE high” if the ER level of the individual exceeds a predetermined value, and “ER/AGE low” if the ER level of the individual does not exceed said predetermined value. In a specific embodiment of this method, said predetermined value of ER is calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said individual. In another specific embodiment, said individual is ER⁻, *BRCA1*, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 1. In another specific embodiment, said individual is ER⁻, sporadic, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 2. In another specific embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 3. In another specific embodiment, said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 5. In another specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

[00125] Where information is available regarding the LN status of a breast cancer patient, the patient may be identified as having a “very good prognosis,” an “intermediate prognosis,” or a poor prognosis, which enables the refinement of treatment. In one embodiment, the invention provides a method of assigning a therapeutic regimen to a breast cancer patient, comprising: (a) classifying said patient as having a “poor prognosis,” “intermediate prognosis,” or “very good prognosis” on the basis of the levels of expression of at least five genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5; and (b) assigning said patient a therapeutic regimen, said therapeutic regimen (i) comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a

good prognosis or an intermediate prognosis, or (ii) comprising chemotherapy if said patient has any other combination of lymph node status and expression profile.

[00126] In another embodiment, a breast cancer patient is assigned a prognosis by a method comprising (a) determining the breast cancer patient's age, ER status, LN status and tumor type; (b) classifying said patient as ER⁻, sporadic; ER⁻, *BRCA1*; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (c) determining an expression profile comprising at least five genes in a cell sample taken from said breast cancer patient wherein markers for said at least five genes are listed in Table 1 if said patient is classified as ER⁻, sporadic; Table 2 if said patient is classified as ER⁻, *BRCA1*; Table 3 if said patient is classified as ER⁺, ER/AGE high; Table 4 if said patient is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said patient is classified as ER⁺, ER/AGE high, LN⁻; (d) determining the similarity of the expression profile of said at least five genes to a template profile comprising levels of expression of said at least five genes to obtain a patient similarity value; (e) comparing said patient similarity value to selected first and second threshold values of similarity, respectively, wherein said second similarity threshold indicates greater similarity to said template expression profile than does said first similarity threshold; and (f) classifying said breast cancer patient as having a first prognosis if said patient similarity value exceeds said second threshold similarity values, a second prognosis if said patient similarity value exceeds said first threshold similarity value but does not exceed said second threshold similarity value, and a third prognosis if said patient similarity value does not exceed said first threshold similarity value. In a specific embodiment of the method, said first prognosis is a "very good prognosis," said second prognosis is an "intermediate prognosis," and said third prognosis is a "poor prognosis," wherein said breast cancer patient is assigned a therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and expression profile.

[00127] The invention also provides a method of assigning a therapeutic regimen to a breast cancer patient, comprising: (a) determining the lymph node status for said patient; (b) determining the expression of at least five genes for which markers are listed in Table 5 in a cell sample from said patient, thereby generating an expression profile; (c) classifying said patient as having a "poor prognosis," "intermediate prognosis," or "very good prognosis" on the basis of said expression profile; and (d) assigning said patient a therapeutic regimen, said therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node

negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and classification. In a specific embodiment of this method, said therapeutic regimen assigned to lymph node negative patients classified as having an “intermediate prognosis” additionally comprises adjuvant hormonal therapy. In another specific embodiment of this method, said classifying step (c) is carried out by a method comprising: (a) rank ordering in descending order a plurality of breast cancer tumor samples that compose a pool of breast cancer tumor samples by the degree of similarity between the expression profile of said at least five genes in each of said tumor samples and the expression profile of said at least five genes across all remaining tumor samples that compose said pool, said degree of similarity being expressed as a similarity value; (b) determining an acceptable number of false negatives in said classifying step, wherein a false negative is a breast cancer patient for whom the expression levels of said at least five genes in said cell sample predicts that said breast cancer patient will have no distant metastases within the first five years after initial diagnosis, but who has had a distant metastasis within the first five years after initial diagnosis; (c) determining a similarity value above which in said rank ordered list said acceptable number of tumor samples or fewer are false negatives; (d) selecting said similarity value determined in step (c) as a first threshold similarity value; (e) selecting a second similarity value, greater than said first similarity value, as a second threshold similarity value; and (f) determining the similarity between the expression profile of said at least five genes in a breast cancer tumor sample from the breast cancer patient and the expression profile of said respective at least five genes in said pool, to obtain a patient similarity value, wherein if said patient similarity value equals or exceeds said second threshold similarity value, said patient is classified as having a “very good prognosis”; if said patient similarity value equals or exceeds said first threshold similarity value, but is less than said second threshold similarity value, said patient is classified as having an “intermediate prognosis”; and if said patient similarity value is less than said first threshold similarity value, said patient is classified as having a “poor prognosis.” Another specific embodiment of this method comprises determining the estrogen receptor (ER) status of said patient, wherein if said patient is ER positive and lymph node negative, said therapeutic regimen assigned to said patient additionally comprises adjuvant hormonal therapy.

[00128] A patient in any patient subset or clinical class, e.g., any one of the classes described above, can be classified as having a particular prognosis level, e.g., a good prognosis or a poor prognosis, based on the similarity of the patient’s cellular constituent profile to an

appropriate template profile for the prognosis level of patients in the clinical class. In one embodiment, a cellular constituent profile corresponding to a certain prognosis level, e.g., a profile comprising measurements of the plurality of cellular constituents representative of levels of the cellular constituents in a plurality of patients having the prognosis level is used as a template for the prognosis level. For example, a good prognosis template profile comprising measurements of the plurality of cellular constituents representative of levels of the cellular constituents in a plurality of good outcome patients or a poor prognosis template profile comprising measurements of the plurality of cellular constituents representative of levels of the cellular constituents in a plurality of poor outcome patients, can be used for determining whether a patient have good or poor prognosis. Here, a good outcome patient is a patient who has non-reoccurrence of metastases within a period of time after initial diagnosis, e.g., a period of 1, 2, 3, 4, 5 or 10 years. In contrast, a poor outcome patient is a patient who has reoccurrence of metastases within a period of time after initial diagnosis, e.g., a period of 1, 2, 3, 4, 5 or 10 years. In a preferred embodiment, both periods are 10 years. Tables 1-5 show exemplary template profiles for the respective patient classes. For example, the expression profile of a patient with a combination of ER+, ER/AGE low, LN+ can be compared with the good prognosis template of Table 4 to determine if the patient has good prognosis or poor prognosis.

[00129] The degree of similarity of the patient's cellular constituent profile to a template of a particular prognosis can be used to indicate whether the patient has the particular prognosis. For example, a high degree of similarity indicates that the patient has the particular prognosis, whereas a low degree of similarity indicates that the patient does not have the particular prognosis. In a preferred embodiment, a patient is classified as having a good prognosis profile if the patient's cellular constituent profile has a high similarity to a good prognosis template and/or has a low similarity to a poor prognosis template. In another embodiment, a patient is classified as having a poor prognosis profile if the patient's cellular constituent profile has a low similarity to a good prognosis template and/or has a high similarity to a poor prognosis template. In embodiments for predicting the responsiveness of a breast cancer patient under the age of 55, the patients in the good and poor outcome patient populations used to generate the templates are preferably also under the age of 55 at the time of diagnosis of breast cancer.

[00130] The degree of similarity between a patient's cellular constituent profile and a template profile can be determined using any method known in the art. In one embodiment, the similarity is represented by a correlation coefficient between the patient's profile and the

template. In one embodiment, a correlation coefficient above a correlation threshold indicates high similarity, whereas a correlation coefficient below the threshold indicates low similarity. In preferred embodiments, the correlation threshold is set as 0.3, 0.4, 0.5 or 0.6. In another embodiment, similarity between a patient's profile and a template is represented by a distance between the patient's profile and the template. In one embodiment, a distance below a given value indicates high similarity, whereas a distance equal to or greater than the given value indicates low similarity.

[00131] As an illustration, in one embodiment, a template for a good prognosis is defined as \bar{z}_1 (e.g., a profile consisting of the xdev's listed in the good prognosis column of one of Tables 1-5) and/or a template for poor prognosis is defined as \bar{z}_2 (e.g., a profile consisting of the xdev's listed in the poor prognosis column of one of Tables 1-5). Either one or both of the two classifier parameters (P_1 and P_2) can then be used to measure degrees of similarities between a patient's profile and the respective templates: P_1 measures the similarity between the patient's profile \bar{y} and the good prognosis template \bar{z}_1 , and P_2 measures the similarity between \bar{y} and the poor prognosis template \bar{z}_2 . In embodiments which employ correlation coefficients, the correlation coefficient P_i can be calculated as:

$$P_i = (\bar{z}_i \bullet \bar{y}) / (\|\bar{z}_i\| \cdot \|\bar{y}\|) \quad (4)$$

where $i = 1$ and 2 .

[00132] Thus, in one embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a selected correlation threshold or if P_2 is equal to or less than a selected correlation threshold. In another embodiment, \bar{y} is classified as a poor prognosis profile if P_1 is less than a selected correlation threshold or if P_2 is above a selected correlation threshold. In still another embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a first selected correlation threshold and \bar{y} is classified as a poor prognosis profile if P_2 is greater than a second selected correlation threshold.

[00133] In a preferred embodiment, the cellular constituent profile is an expression profile comprising measurements of a plurality of transcripts (e.g., measured as mRNAs or cDNAs) in a sample derived from a patient, e.g., the plurality of transcripts corresponding to the markers in all or a portion of one of Tables 1-5. In this embodiment, the good prognosis template can be a good prognosis expression template comprising measurements of the

plurality of transcripts representative of expression levels of the transcripts in a plurality of good prognosis patients, and the poor prognosis template can be a poor prognosis expression template comprising measurements of the plurality of transcripts representative of expression levels of the transcripts in a plurality of poor prognosis patients. In a preferred embodiment, measurement of each transcript in the good or poor prognosis expression template is an average of expression levels of the transcript in the plurality of good or poor prognosis patients, respectively.

[00134] In another embodiment, the expression profile is a differential expression profile comprising differential measurements of the plurality of transcripts in a sample derived from the patient versus measurements of the plurality of transcripts in a control sample. The differential measurements can be x_{dev} , $\log(\text{ratio})$, error-weighted $\log(\text{ratio})$, or a mean subtracted $\log(\text{intensity})$ (see, e.g., Stoughton et al., PCT publication WO 00/39339, published on July 6, 2000; U.S. Patent Application No. 10/848,755, filed May 18, 2004, by Mao et al., attorney docket no: 9301-188-999, each of which is incorporated herein by reference in its entirety).

5.4.3 IMPROVING SENSITIVITY TO EXPRESSION LEVEL DIFFERENCES

[00135] In using the markers disclosed herein, and, indeed, using any sets of markers, e.g., to compare profiles or to differentiate an individual having one phenotype from another individual having a second phenotype, one can compare the profile comprising absolute expression levels of the markers in a sample to a template; for example, a template comprising the average levels of expression of the markers in a plurality of individuals. To increase the sensitivity of the comparison, however, the expression level values are preferably transformed in a number of ways. Also, to differentiate an individual having one phenotype from another individual having a second phenotype using any sets of markers, one can compare the absolute expression of each of the markers in a sample to a control; for example, the control can be the average level of expression of each of the markers, respectively, in a pool of individuals.

[00136] For example, the expression level of each of the markers can be normalized by the average expression level of all markers the expression level of which is determined, or by the average expression level of a set of control genes. Thus, in one embodiment, the markers are represented by probes on a microarray, and the expression level of each of the markers is normalized by the mean or median expression level across all of the genes represented on the microarray, including any non-marker genes. In a specific embodiment, the normalization is

carried out by dividing the median or mean level of expression of all of the genes on the microarray. In another embodiment, the expression levels of the markers is normalized by the mean or median level of expression of a set of control markers. In a specific embodiment, the control markers comprise a set of housekeeping genes. In another specific embodiment, the normalization is accomplished by dividing by the median or mean expression level of the control genes.

[00137] The sensitivity of a marker-based assay will also be increased if the expression levels of individual markers are compared to the expression of the same markers in a pool of samples. Preferably, the comparison is to the mean or median expression level of each the marker genes in the pool of samples. Such a comparison may be accomplished, for example, by dividing by the mean or median expression level of the pool for each of the markers from the expression level each of the markers in the sample. This has the effect of accentuating the relative differences in expression between markers in the sample and markers in the pool as a whole, making comparisons more sensitive and more likely to produce meaningful results than the use of absolute expression levels alone. The expression level data may be transformed in any convenient way; preferably, the expression level data for all is log transformed before means or medians are taken.

[00138] In performing comparisons to a pool, two approaches may be used. First, the expression levels of the markers in the sample may be compared to the expression level of those markers in the pool, where nucleic acid derived from the sample and nucleic acid derived from the pool are hybridized during the course of a single experiment. Such an approach requires that new pool nucleic acid be generated for each comparison or limited numbers of comparisons, and is therefore limited by the amount of nucleic acid available. Alternatively, and preferably, the expression levels in a pool, whether normalized and/or transformed or not, are stored on a computer, or on computer-readable media, to be used in comparisons to the individual expression level data from the sample (i.e., single-channel data).

[00139] The current invention also provides the following method of classifying a first cell or organism as having one of at least two different phenotypes, where the different phenotypes comprise a first phenotype and a second phenotype. The level of expression of each of a plurality of markers in a first sample from the first cell or organism is compared to the level of expression of each of said markers, respectively, in a pooled sample from a plurality of cells or organisms, the plurality of cells or organisms comprising different cells or organisms exhibiting said at least two different phenotypes, respectively, to produce a first compared

value. The first compared value is then compared to a second compared value, wherein said second compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having said first phenotype to the level of expression of each of said markers, respectively, in the pooled sample. The first compared value is then compared to a third compared value, wherein said third compared value is the product of a method comprising comparing the level of expression of each of the markers in a sample from a cell or organism characterized as having the second phenotype to the level of expression of each of the markers, respectively, in the pooled sample. In specific embodiments, the marker can be a gene, a protein encoded by the gene, etc. Optionally, the first compared value can be compared to additional compared values, respectively, where each additional compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having a phenotype different from said first and second phenotypes but included among the at least two different phenotypes, to the level of expression of each of said genes, respectively, in said pooled sample. Finally, a determination is made as to which of said second, third, and, if present, one or more additional compared values, said first compared value is most similar, wherein the first cell or organism is determined to have the phenotype of the cell or organism used to produce said compared value most similar to said first compared value.

[00140] The sensitivity of a marker-based assay will also be increased if the expression levels of individual markers are compared to the expression of the same markers in a control sample, e.g., a sample comprises a pool of samples, to generate a differential expression profile. Such a comparison may be accomplished, for example, by determining a ratio between expression level of each marker in the sample and the expression level of the corresponding marker in the control sample. This has the effect of accentuating the relative differences in expression between markers in the sample and markers in the control as a whole, making subsequent comparisons to a template more sensitive and more likely to produce meaningful results than the use of absolute expression levels alone. The comparison may be performed in any convenient way, e.g., by taking difference, ratio, or log(ratio).

[00141] In performing comparisons to a control sample, two approaches may be used. First, the expression levels of the markers in the sample may be compared to the expression level of those markers in the control sample, where nucleic acid derived from the sample and nucleic acid derived from the control are hybridized during the course of a single experiment. Such an approach requires that new control sample of nucleic acid be generated for each

comparison or limited numbers of comparisons, and is therefore limited by the amount of nucleic acid available. Alternatively, the expression levels in a control sample, whether normalized and/or transformed or not, are stored on a computer, or on computer-readable media, to be used in comparisons to the individual expression level data from the sample (i.e., single-channel data).

[00142] The methods of the invention preferably use a control or reference sample, which can be any suitable sample against which changes in cellular constituents can be determined. In one embodiment, the control or reference sample is generated by pooling together the plurality of cellular constituents, e.g., a plurality of transcripts or cDNAs, or a plurality of protein species, from a plurality of breast cancer patients. Alternatively, the control or reference sample can be generated by pooling together purified or synthesized cellular constituents, e.g., a plurality of purified or synthesized transcripts or cDNAs, a plurality of purified or synthesized protein species. In one embodiment, synthetic RNAs for each transcripts or cDNAs are pooled to form the control or reference sample. Preferably, the abundances of synthetic RNAs are approximately the abundances of the corresponding transcripts in a real tumor pool. The differential expression of marker genes for each individual patient sample is measured against this control sample. In one embodiment, 60-mer oligonucleotides corresponding to the probe sequences on a microarray used to assay the expression levels of the diagnostic/prognostic transcripts are synthesized and cloned into pBluescript SK- vector (Statagene, La Jolla, CA), adjacent to the T7 promotor sequence. Individual clones are isolated, and the sequences of their inserts are verified by DNA sequencing. To generate synthetic RNAs, clones are linearized with *EcoRI* and a T7 in vitro transcription (IVT) reaction is performed by MegaScript kit (Ambion, Austin, TX), followed by DNase treatment of the product. Synthetic RNAs are purified on RNeasy columns (Qiagen, Valencia, CA). These synthetic RNAs are transcribed, amplified, labeled, and mixed together to make the reference pool. The abundance of those synthetic RNAs are chosen to approximate the abundances of the transcripts of the corresponding marker genes in the real tumor pool.

[00143] The current invention provides the following method of classifying a first cell or organism as having one of at least two different phenotypes, where the different phenotypes comprise a first phenotype and a second phenotype. The level of expression of each of a plurality of markers in a first sample from the first cell or organism is compared to the level of expression of each of said markers, respectively, in a pooled sample from a plurality of cells or organisms, the plurality of cells or organisms comprising different cells or organisms

exhibiting said at least two different phenotypes, respectively, to produce a first compared value so that a first differential profile comprising a plurality of first compared values for said plurality of markers is generated. The first differential profile is then compared to a second differential profile comprising a plurality of second compared values, wherein each said second compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having said first phenotype to the level of expression of each of said markers, respectively, in the pooled sample. The first differential profile is then compared to a third differential profile comprising a plurality of third compared values, wherein each said third compared value is the product of a method comprising comparing the level of expression of each of the markers in a sample from a cell or organism characterized as having the second phenotype to the level of expression of each of the markers, respectively, in the pooled sample. In specific embodiments, each marker can be a gene, a protein encoded by the gene, etc. Optionally, the first differential profile can be compared to additional expression profiles each of which comprising additional compared values, respectively, where each additional compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having a phenotype different from said first and second phenotypes but included among the at least two different phenotypes, to the level of expression of each of said genes, respectively, in said pooled sample. Finally, a determination is made as to which of said second, third, and, if present, one or more additional differential profiles, said first differential profile is most similar, wherein the first cell or organism is determined to have the phenotype of the cell or organism used to produce said differential profile most similar to said first differential profile.

[00144] In a specific embodiment of this method, the compared values are each ratios of the levels of expression of each of said genes. In another specific embodiment, each of the levels of expression of each of the genes in the pooled sample are normalized prior to any of the comparing steps. In a more specific embodiment, the normalization of the levels of expression is carried out by dividing by the median or mean level of the expression of each of the genes or dividing by the mean or median level of expression of one or more housekeeping genes in the pooled sample from said cell or organism. In another specific embodiment, the normalized levels of expression are subjected to a log transform, and the comparing steps comprise subtracting the log transform from the log of the levels of expression of each of the genes in the sample. In another specific embodiment, the two or more different phenotypes are different stages of a disease or disorder. In still another specific embodiment, the two or

more different phenotypes are different prognoses of a disease or disorder. In yet another specific embodiment, the levels of expression of each of the genes, respectively, in the pooled sample or said levels of expression of each of said genes in a sample from the cell or organism characterized as having the first phenotype, second phenotype, or said phenotype different from said first and second phenotypes, respectively, are stored on a computer or on a computer-readable medium.

[00145] In another specific embodiment, the two phenotypes are good prognosis and poor prognosis. In a more specific embodiment, the two phenotypes are good prognosis and poor prognosis for an individual that is identified as having ER⁻, *BRCA1* status, ER⁻, sporadic status, ER⁺, ER/AGE high status, ER⁺, ER/AGE low, LN⁺ status, or ER⁺, ER/AGE low, LN⁺ status.

[00146] In another specific embodiment, the comparison is made between the expression profile of the genes in the sample and the expression profile of the same genes in a pool representing only one of two or more phenotypes. In the context of prognosis-correlated genes, for example, one can compare the expression levels of prognosis-related genes in a sample to the average levels of the expression of the same genes in a plurality of “good prognosis” samples (as opposed to a plurality of samples that include samples from patients having poor prognoses and good prognoses). Thus, in this method, a sample is classified as having a good prognosis if the expression profile of prognosis-correlated genes exceeds a chosen coefficient of correlation to the average “good prognosis” expression profile (*e.g.*, the profile comprising average levels of expression of prognosis-correlated genes in samples from a plurality of patients having a “good prognosis”). Patients whose expression profiles correlate more poorly with the “good prognosis” expression profile (*e.g.*, whose correlation coefficient fails to exceed the chosen coefficient) are classified as having a poor prognosis.

[00147] Where individuals are classified on the basis of phenotypic, genotypic, or clinical characteristics into patient subsets, the pool of samples may be a pool of samples for the phenotype that includes samples representing each of the patient subsets. Alternatively, the pool of samples may be a pool of samples for the phenotype representing only the specific patient subset. For example, where an individual is classified as ER⁺, sporadic, the pool of samples to which the individual’s sample is compared may be a pool of samples from ER⁺, sporadic individuals having a good prognosis only, or may be a pool of samples of individuals having a good prognosis, without regard to ER status or mutation type.

[00148] The method can be applied to a plurality of patient subsets. For example, in a specific embodiment, the phenotype is good prognosis, and the individual is classified into

one of the following patient subsets: ER⁻, *BRCA1* status, ER⁻, sporadic status, ER⁺, ER/AGE high status, ER⁺, ER/AGE low, LN⁺ status, or ER⁺, ER/AGE low, LN⁺ status. A set of markers informative for prognosis for the patient subset into which the individual is classified is then used to determine the likely prognosis for the individual. A sample is classified as coming from an individual having a good prognosis if the expression profile of prognosis-correlated genes for the particular subset into which the individual is classified exceeds a chosen coefficient of correlation to the average “good prognosis” expression profile (*e.g.*, the levels of expression of prognosis-correlated genes in a plurality of samples from patients within the subclass having a “good prognosis”). Patients whose expression levels correlate more poorly with the “good prognosis” expression profile (*e.g.*, whose correlation coefficient fails to exceed the chosen coefficient) are classified as having a poor prognosis.

[00149] Of course, single-channel data may also be used without specific comparison to a mathematical sample pool. For example, a sample may be classified as having a first or a second phenotype, wherein the first and second phenotypes are related, by calculating the similarity between the expression profile of at least 5 markers in the sample, where the markers are correlated with the first or second phenotype, to a first phenotype template and a second phenotype template each comprising the expression levels of the same markers, by (a) labeling nucleic acids derived from a sample with a fluorophore to obtain a pool of fluorophore-labeled nucleic acids; (b) contacting said fluorophore-labeled nucleic acid with a microarray under conditions such that hybridization can occur, detecting at each of a plurality of discrete loci on the microarray a fluorescent emission signal from said fluorophore-labeled nucleic acid that is bound to said microarray under said conditions; and (c) determining the similarity of marker gene expression in the individual sample to the first and second templates, wherein if said expression is more similar to the first template, the sample is classified as having the first phenotype, and if said expression is more similar to the second template, the sample is classified as having the second phenotype.

[0100] In a specific embodiment of the above method, the first phenotype is a good prognosis of breast cancer, the sample is a sample from an individual that has been classified into a patient subset, and the first and second templates are templates for the phenotype for the particular patient subset. In a more specific embodiment, for example, the first phenotype is a good prognosis, the second phenotype is a poor prognosis, the patient is classified into an ER⁻, sporadic patient subset, an ER⁻, *BRCA1* subset, an ER⁺, ER/AGE high subset, an ER⁺, ER/AGE low, LN⁺ subset, or an ER⁺, ER/AGE low, LN⁺ subset, and said first and second

templates are templates derived from the expression of the marker genes in individuals having a good prognosis and a poor prognosis, respectively, wherein said individuals are all of the patient subset into which said patient is classified.

5.5 DETERMINATION OF MARKER GENE EXPRESSION LEVELS

5.5.1 METHODS

[00150] The expression levels of the marker genes in a sample may be determined by any means known in the art. The expression level may be determined by isolating and determining the level (*i.e.*, amount) of nucleic acid transcribed from each marker gene. Alternatively, or additionally, the level of specific proteins encoded by a marker gene may be determined.

[00151] The level of expression of specific marker genes can be accomplished by determining the amount of mRNA, or polynucleotides derived therefrom, present in a sample. Any method for determining RNA levels can be used. For example, RNA is isolated from a sample and separated on an agarose gel. The separated RNA is then transferred to a solid support, such as a filter. Nucleic acid probes representing one or more markers are then hybridized to the filter by northern hybridization, and the amount of marker-derived RNA is determined. Such determination can be visual, or machine-aided, for example, by use of a densitometer. Another method of determining RNA levels is by use of a dot-blot or a slot-blot. In this method, RNA, or nucleic acid derived therefrom, from a sample is labeled. The RNA or nucleic acid derived therefrom is then hybridized to a filter containing oligonucleotides derived from one or more marker genes, wherein the oligonucleotides are placed upon the filter at discrete, easily-identifiable locations. Hybridization, or lack thereof, of the labeled RNA to the filter-bound oligonucleotides is determined visually or by densitometer. Polynucleotides can be labeled using a radiolabel or a fluorescent (*i.e.*, visible) label.

[00152] These examples are not intended to be limiting; other methods of determining RNA abundance are known in the art.

[00153] The level of expression of particular marker genes may also be assessed by determining the level of the specific protein expressed from the marker genes. This can be accomplished, for example, by separation of proteins from a sample on a polyacrylamide gel, followed by identification of specific marker-derived proteins using antibodies in a western blot. Alternatively, proteins can be separated by two-dimensional gel electrophoresis

systems. Two-dimensional gel electrophoresis is well-known in the art and typically involves isoelectric focusing along a first dimension followed by SDS-PAGE electrophoresis along a second dimension. *See, e.g.*, Hames *et al.*, 1990, GEL ELECTROPHORESIS OF PROTEINS: A PRACTICAL APPROACH, IRL Press, New York; Shevchenko *et al.*, *Proc. Nat'l Acad. Sci. USA* 93:1440-1445 (1996); Sagliocco *et al.*, *Yeast* 12:1519-1533 (1996); Lander, *Science* 274:536-539 (1996). The resulting electropherograms can be analyzed by numerous techniques, including mass spectrometric techniques, western blotting and immunoblot analysis using polyclonal and monoclonal antibodies.

[00154] Alternatively, marker-derived protein levels can be determined by constructing an antibody microarray in which binding sites comprise immobilized, preferably monoclonal, antibodies specific to a plurality of protein species encoded by the cell genome. Preferably, antibodies are present for a substantial fraction of the marker-derived proteins of interest. Methods for making monoclonal antibodies are well known (*see, e.g.*, Harlow and Lane, 1988, ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor, New York, which is incorporated in its entirety for all purposes). In one embodiment, monoclonal antibodies are raised against synthetic peptide fragments designed based on genomic sequence of the cell. With such an antibody array, proteins from the cell are contacted to the array, and their binding is assayed with assays known in the art. Generally, the expression, and the level of expression, of proteins of diagnostic or prognostic interest can be detected through immunohistochemical staining of tissue slices or sections.

[00155] Finally, expression of marker genes in a number of tissue specimens may be characterized using a "tissue array" (Kononen *et al.*, *Nat. Med* 4(7):844-7 (1998)). In a tissue array, multiple tissue samples are assessed on the same microarray. The arrays allow *in situ* detection of RNA and protein levels; consecutive sections allow the analysis of multiple samples simultaneously.

5.5.2 MICROARRAYS

[00156] In preferred embodiments, polynucleotide microarrays are used to measure expression so that the expression status of each of the markers above is assessed simultaneously. Generally, microarrays according to the invention comprise a plurality of markers informative for prognosis, or outcome determination, for a particular disease or condition, and, in particular, for individuals having specific combinations of genotypic or phenotypic characteristics of the disease or condition (*i.e.*, that are prognosis-informative for a particular patient subset).

[00157] The microarrays of the invention preferably comprise at least 2, 3, 4, 5, 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200 or more of markers, or all of the markers, or any combination of markers, identified as prognosis-informative within a patient subset. The actual number of informative markers the microarray comprises will vary depending upon the particular condition of interest, the number of markers identified, and, optionally, the number of informative markers found to result in the least Type I error, Type II error, or Type I and Type II error in determination of prognosis. As used herein, "Type I error" means a false positive and "Type II error" means a false negative; in the example of prognosis of breast cancer, Type I error is the mis-characterization of an individual with a good prognosis as having a poor prognosis, and Type II error is the mis-characterization of an individual with a poor prognosis as having a good prognosis.

[00158] In specific embodiments, the invention provides polynucleotide arrays in which the prognosis markers identified for a particular patient subset comprise at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the probes on said array. In another specific embodiment, the microarray comprises a plurality of probes, wherein said plurality of probes comprise probes complementary and hybridizable to at least 75% of the prognosis-informative markers identified for a particular patient subset. Microarrays of the invention, of course, may comprise probes complementary and hybridizable to prognosis-informative markers for a plurality of the patient subsets, or for each patient subset, identified for a particular condition. In another embodiment, therefore, the microarray of the invention comprises a plurality of probes complementary and hybridizable to at least 75% of the prognosis-informative markers identified for each patient subset identified for the condition of interest, and wherein said probes, in total, are at least 50% of the probes on said microarray.

[00159] In yet another specific embodiment, microarrays that are used in the methods disclosed herein optionally comprise markers additional to at least some of the markers identified by the methods disclosed elsewhere herein. For example, in a specific embodiment, the microarray is a screening or scanning array as described in Altschuler *et al.*, International Publication WO 02/18646, published March 7, 2002 and Scherer *et al.*, International Publication WO 02/16650, published February 28, 2002. The scanning and screening arrays comprise regularly-spaced, positionally-addressable probes derived from genomic nucleic acid sequence, both expressed and unexpressed. Such arrays may comprise probes corresponding to a subset of, or all of, the markers identified for the patient subset(s) for the condition of interest, and can be used to monitor marker expression in the same way as a microarray containing only prognosis-informative markers otherwise identified.

[00160] In yet another specific embodiment, the microarray is a commercially-available cDNA microarray that comprises at least five markers identified by the methods described herein. Preferably, a commercially-available cDNA microarray comprises all of the markers identified by the methods described herein as being informative for a patient subset for a particular condition. However, such a microarray may comprise at least 5, 10, 15 or 25 of such markers, up to the maximum number of markers identified.

[00161] In an embodiment specific to breast cancer, the invention provides for oligonucleotide or cDNA arrays comprising probes hybridizable to the genes corresponding to each of the marker sets described above (*i.e.*, markers informative for ER⁻, sporadic individuals, markers informative for ER⁻, *BRCAl* individuals, markers informative for ER⁺, ER/AGE high individuals, markers informative for ER⁺, ER/AGE low, LN⁺ individuals, and markers informative for ER⁺, ER/AGE low, LN⁻ individuals, as shown in Tables 1-5). Any of the microarrays described herein may be provided in a sealed container in a kit.

[00162] The invention provides microarrays containing probes useful for the prognosis of any breast cancer patient, or for breast cancer patients classified into one of a plurality of patient subsets. In particular, the invention provides polynucleotide arrays comprising probes to a subset or subsets of at least 5, 10, 15, 20, 25 or more of the genetic markers, or up to the full set of markers, in any of Tables 1-5, which distinguish between patients with good and poor prognosis. In certain embodiments, therefore, the invention provides microarrays comprising probes for a plurality of the genes for which markers are listed in Tables 1, 2, 3, 4 or 5. In a specific embodiment, the microarray of the invention comprises 1, 2, 3, 4, 5 or 10 of the markers in Table 1, at least five of the markers in Table 2; 1, 2, 3, 4, 5 or 10 of the markers in Table 3; 1, 2, 3, 4, 5 or 10 of the markers in Table 4; or 1, 2, 3, 4, 5 or 10 of the markers in Table 1. In other embodiments, the microarray comprises probes for 1, 2, 3, 4, 5, or 10 of the markers shown in any two, three or four of Tables 1-5, or all of Tables 1-5. In other embodiments, the microarray of the invention contains each of the markers in Table 1, Table 2, Table 3, Table 4, or Table 5. In another embodiment, the microarray contains all of the markers shown in Tables 1-5. In specific embodiments, the array comprises probes derived only from the markers listed in Table 1, Table 2, Table 3, Table 4, or Table 5; probes derived from any two of Tables 1-5; any three of Tables 1-5; any four of Tables 1-5; or all of Tables 1-5.

[00163] In other embodiments, the array comprises a plurality of probes derived from markers listed in any of Tables 1-5 in combination with a plurality of other probes, derived

from markers not listed in any of Tables 1-5, that are identified as informative for the prognosis of breast cancer.

[00164] In specific embodiments, the invention provides polynucleotide arrays in which the breast cancer prognosis markers described herein in Tables 1, 2, 3, 4 and/or 5 comprise at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the probes on said array. In another specific embodiment, the microarray comprises a plurality of probes, wherein said plurality of probes comprise probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 1; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 2; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 3; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 4; and probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 5, wherein said probes, in total, are at least 50% of the probes on said microarray.

[00165] In yet another specific embodiment, microarrays that are used in the methods disclosed herein optionally comprise markers additional to at least some of the markers listed in Tables 1-5. For example, in a specific embodiment, the microarray is a screening or scanning array as described in Altschuler *et al.*, International Publication WO 02/18646, published March 7, 2002 and Scherer *et al.*, International Publication WO 02/16650, published February 28, 2002. The scanning and screening arrays comprise regularly-spaced, positionally-addressable probes derived from genomic nucleic acid sequence, both expressed and unexpressed. Such arrays may comprise probes corresponding to a subset of, or all of, the markers listed in Tables 1-5, or a subset thereof as described above, and can be used to monitor marker expression in the same way as a microarray containing only markers listed in Tables 1-5.

[00166] In yet another specific embodiment, the microarray is a commercially-available cDNA microarray that comprises at least five of the markers listed in Tables 1-5. Preferably, a commercially-available cDNA microarray comprises all of the markers listed in Tables 1-5. However, such a microarray may comprise at least 5, 10, 15 or 25 of the markers in any of Tables 1-5, up to the maximum number of markers in a Table, and may comprise all of the markers in any one of Tables 1-5, and a subset of another of Tables 1-5, or subsets of each as described above. In a specific embodiment of the microarrays used in the methods disclosed herein, the markers that are all or a portion of Tables 1-5 make up at least 50%, 60%, 70%, 80%, 90%, 95% or 98% of the probes on the microarray.

[00167] General methods pertaining to the construction of microarrays comprising the marker sets and/or subsets above are described in the following sections.

[00168]

[00169]

5.5.2.1 CONSTRUCTION OF MICROARRAYS

[00170] Microarrays are prepared by selecting probes which comprise a polynucleotide sequence, and then immobilizing such probes to a solid support or surface. For example, the probes may comprise DNA sequences, RNA sequences, or copolymer sequences of DNA and RNA. The polynucleotide sequences of the probes may also comprise DNA and/or RNA analogues, or combinations thereof. For example, the polynucleotide sequences of the probes may be full or partial fragments of genomic DNA. The polynucleotide sequences of the probes may also be synthesized nucleotide sequences, such as synthetic oligonucleotide sequences. The probe sequences can be synthesized either enzymatically *in vivo*, enzymatically *in vitro* (e.g., by PCR), or non-enzymatically *in vitro*.

[00171] The probe or probes used in the methods of the invention are preferably immobilized to a solid support which may be either porous or non-porous. For example, the probes of the invention may be polynucleotide sequences which are attached to a nitrocellulose or nylon membrane or filter covalently at either the 3' or the 5' end of the polynucleotide. Such hybridization probes are well known in the art (see, e.g., Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). Alternatively, the solid support or surface may be a glass or plastic surface. In a particularly preferred embodiment, hybridization levels are measured to microarrays of probes consisting of a solid phase on the surface of which are immobilized a population of polynucleotides, such as a population of DNA or DNA mimics, or, alternatively, a population of RNA or RNA mimics. The solid phase may be a nonporous or, optionally, a porous material such as a gel.

[00172] In preferred embodiments, a microarray comprises a support or surface with an ordered array of binding (e.g., hybridization) sites or "probes" each representing one of the markers described herein. Preferably the microarrays are addressable arrays, and more preferably positionally addressable arrays. More specifically, each probe of the array is preferably located at a known, predetermined position on the solid support such that the identity (*i.e.*, the sequence) of each probe can be determined from its position in the array

(*i.e.*, on the support or surface). In preferred embodiments, each probe is covalently attached to the solid support at a single site.

[00173] Microarrays can be made in a number of ways, of which several are described below. However produced, microarrays share certain characteristics. The arrays are reproducible, allowing multiple copies of a given array to be produced and easily compared with each other. Preferably, microarrays are made from materials that are stable under binding (*e.g.*, nucleic acid hybridization) conditions. The microarrays are preferably small, *e.g.*, between 1 cm² and 25 cm², between 12 cm² and 13 cm², or 3 cm². However, larger arrays are also contemplated and may be preferable, *e.g.*, for use in screening arrays. Preferably, a given binding site or unique set of binding sites in the microarray will specifically bind (*e.g.*, hybridize) to the product of a single gene in a cell (*e.g.*, to a specific mRNA, or to a specific cDNA derived therefrom). However, in general, other related or similar sequences will cross hybridize to a given binding site.

[00174] The microarrays of the present invention include one or more test probes, each of which has a polynucleotide sequence that is complementary to a subsequence of RNA or DNA to be detected. Preferably, the position of each probe on the solid surface is known. Indeed, the microarrays are preferably positionally addressable arrays. Specifically, each probe of the array is preferably located at a known, predetermined position on the solid support such that the identity (*i.e.*, the sequence) of each probe can be determined from its position on the array (*i.e.*, on the support or surface).

[00175] According to the invention, the microarray is an array (*i.e.*, a matrix) in which each position represents one of the markers described herein. For example, each position can contain a DNA or DNA analogue based on genomic DNA to which a particular RNA or cDNA transcribed from that genetic marker can specifically hybridize. The DNA or DNA analogue can be, *e.g.*, a synthetic oligomer or a gene fragment. In one embodiment, probes representing each of the markers is present on the array. In a preferred embodiment, the array comprises probes for each of the markers listed in Tables 1-5.

5.5.2.2 PREPARING PROBES FOR MICROARRAYS

[00176] As noted above, the “probe” to which a particular polynucleotide molecule specifically hybridizes according to the invention contains a complementary genomic polynucleotide sequence. The probes of the microarray preferably consist of nucleotide sequences of no more than 1,000 nucleotides. In some embodiments, the probes of the array consist of nucleotide sequences of 10 to 1,000 nucleotides. In a preferred embodiment, the

nucleotide sequences of the probes are in the range of 10-200 nucleotides in length and are genomic sequences of a species of organism, such that a plurality of different probes is present, with sequences complementary and thus capable of hybridizing to the genome of such a species of organism, sequentially tiled across all or a portion of such genome. In other specific embodiments, the probes are in the range of 10-30 nucleotides in length, in the range of 10-40 nucleotides in length, in the range of 20-50 nucleotides in length, in the range of 40-80 nucleotides in length, in the range of 50-150 nucleotides in length, in the range of 80-120 nucleotides in length, and most preferably are 60 nucleotides in length.

[00177] The probes may comprise DNA or DNA “mimics” (*e.g.*, derivatives and analogues) corresponding to a portion of an organism’s genome. In another embodiment, the probes of the microarray are complementary RNA or RNA mimics. DNA mimics are polymers composed of subunits capable of specific, Watson-Crick-like hybridization with DNA, or of specific hybridization with RNA. The nucleic acids can be modified at the base moiety, at the sugar moiety, or at the phosphate backbone. Exemplary DNA mimics include, *e.g.*, phosphorothioates.

[00178] DNA can be obtained, *e.g.*, by polymerase chain reaction (PCR) amplification of genomic DNA or cloned sequences. PCR primers are preferably chosen based on a known sequence of the genome that will result in amplification of specific fragments of genomic DNA. Computer programs that are well known in the art are useful in the design of primers with the required specificity and optimal amplification properties, such as *Oligo* version 5.0 (National Biosciences). Typically each probe on the microarray will be between 10 bases and 50,000 bases, usually between 300 bases and 1,000 bases in length. PCR methods are well known in the art, and are described, for example, in Innis *et al.*, eds., PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS, Academic Press Inc., San Diego, CA (1990). It will be apparent to one skilled in the art that controlled robotic systems are useful for isolating and amplifying nucleic acids.

[00179] An alternative, preferred means for generating the polynucleotide probes of the microarray is by synthesis of synthetic polynucleotides or oligonucleotides, *e.g.*, using N-phosphonate or phosphoramidite chemistries (Froehler *et al.*, *Nucleic Acid Res.* 14:5399-5407 (1986); McBride *et al.*, *Tetrahedron Lett.* 24:246-248 (1983)). Synthetic sequences are typically between about 10 and about 500 bases in length, more typically between about 20 and about 100 bases, and most preferably between about 40 and about 70 bases in length. In some embodiments, synthetic nucleic acids include non-natural bases, such as, but by no means limited to, inosine. As noted above, nucleic acid analogues may be used as binding

sites for hybridization. An example of a suitable nucleic acid analogue is peptide nucleic acid (*see, e.g.,* Egholm *et al.*, *Nature* 363:566-568 (1993); U.S. Patent No. 5,539,083).

[00180] Probes are preferably selected using an algorithm that takes into account binding energies, base composition, sequence complexity, cross-hybridization binding energies, and secondary structure. *See* Friend *et al.*, International Patent Publication WO 01/05935, published January 25, 2001; Hughes *et al.*, *Nat. Biotech.* 19:342-7 (2001).

[00181] A skilled artisan will also appreciate that positive control probes, *e.g.*, probes known to be complementary and hybridizable to sequences in the target polynucleotide molecules, and negative control probes, *e.g.*, probes known to not be complementary and hybridizable to sequences in the target polynucleotide molecules, should be included on the array. In one embodiment, positive controls are synthesized along the perimeter of the array. In another embodiment, positive controls are synthesized in diagonal stripes across the array. In still another embodiment, the reverse complement for each probe is synthesized next to the position of the probe to serve as a negative control. In yet another embodiment, sequences from other species of organism are used as negative controls or as "spike-in" controls.

5.5.2.3 ATTACHING PROBES TO THE SOLID SURFACE

[00182] The probes are attached to a solid support or surface, which may be made, *e.g.*, from glass, plastic (*e.g.*, polypropylene, nylon), polyacrylamide, nitrocellulose, gel, or other porous or nonporous material. A preferred method for attaching the nucleic acids to a surface is by printing on glass plates, as is described generally by Schena *et al.*, *Science* 270:467-470 (1995). This method is especially useful for preparing microarrays of cDNA (See also, DeRisi *et al.*, *Nature Genetics* 14:457-460 (1996); Shalon *et al.*, *Genome Res.* 6 :639-645 (1996); and Schena *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:10539-11286 (1995)).

[00183] A second preferred method for making microarrays is by making high-density oligonucleotide arrays. Techniques are known for producing arrays containing thousands of oligonucleotides complementary to defined sequences, at defined locations on a surface using photolithographic techniques for synthesis *in situ* (*see, Fodor et al.*, 1991, *Science* 251:767-773; Pease *et al.*, 1994, *Proc. Natl. Acad. Sci. U.S.A.* 91:5022-5026; Lockhart *et al.*, 1996, *Nature Biotechnology* 14:1675; U.S. Patent Nos. 5,578,832; 5,556,752; and 5,510,270) or other methods for rapid synthesis and deposition of defined oligonucleotides (Blanchard *et al.*, *Biosensors & Bioelectronics* 11:687-690). When these methods are used, oligonucleotides (*e.g.*, 60-mers) of known sequence are synthesized directly on a surface such

as a derivatized glass slide. Usually, the array produced is redundant, with several oligonucleotide molecules per RNA.

[00184] Other methods for making microarrays, *e.g.*, by masking (Maskos and Southern, 1992, *Nuc. Acids. Res.* 20:1679-1684), may also be used. In principle, and as noted *supra*, any type of array, for example, dot blots on a nylon hybridization membrane (see Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)) could be used. However, as will be recognized by those skilled in the art, very small arrays will frequently be preferred because hybridization volumes will be smaller.

[00185] In one embodiment, the arrays of the present invention are prepared by synthesizing polynucleotide probes on a support. In such an embodiment, polynucleotide probes are attached to the support covalently at either the 3' or the 5' end of the polynucleotide.

[00186] In a particularly preferred embodiment, microarrays of the invention are manufactured by means of an ink jet printing device for oligonucleotide synthesis, *e.g.*, using the methods and systems described by Blanchard in U.S. Pat. No. 6,028,189; Blanchard *et al.*, 1996, *Biosensors and Bioelectronics* 11:687-690; Blanchard, 1998, in *Synthetic DNA Arrays in Genetic Engineering*, Vol. 20, J.K. Setlow, Ed., Plenum Press, New York at pages 111-123. Specifically, the oligonucleotide probes in such microarrays are preferably synthesized in arrays, *e.g.*, on a glass slide, by serially depositing individual nucleotide bases in "microdroplets" of a high surface tension solvent such as propylene carbonate. The microdroplets have small volumes (*e.g.*, 100 pL or less, more preferably 50 pL or less) and are separated from each other on the microarray (*e.g.*, by hydrophobic domains) to form circular surface tension wells which define the locations of the array elements (*i.e.*, the different probes). Microarrays manufactured by this ink-jet method are typically of high density, preferably having a density of at least about 2,500 different probes per 1 cm². The polynucleotide probes are attached to the support covalently at either the 3' or the 5' end of the polynucleotide.

5.5.2.4 TARGET POLYNUCLEOTIDE MOLECULES

[00187] The polynucleotide molecules which may be analyzed by the present invention (the "target polynucleotide molecules") may be from any clinically relevant source, but are expressed RNA or a nucleic acid derived therefrom (*e.g.*, cDNA or amplified RNA derived from cDNA that incorporates an RNA polymerase promoter), including naturally occurring nucleic acid molecules, as well as synthetic nucleic acid molecules. In one embodiment, the

target polynucleotide molecules comprise RNA, including, but by no means limited to, total cellular RNA, poly(A)⁺ messenger RNA (mRNA) or fraction thereof, cytoplasmic mRNA, or RNA transcribed from cDNA (*i.e.*, cRNA; see, *e.g.*, Linsley & Schelter, U.S. Patent Application No. 09/411,074, filed October 4, 1999, or U.S. Patent Nos. 5,545,522, 5,891,636, or 5,716,785). Methods for preparing total and poly(A)⁺ RNA are well known in the art, and are described generally, *e.g.*, in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). In one embodiment, RNA is extracted from cells of the various types of interest in this invention using guanidinium thiocyanate lysis followed by CsCl centrifugation (Chirgwin *et al.*, 1979, *Biochemistry* 18:5294-5299). In another embodiment, total RNA is extracted using a silica gel-based column, commercially available examples of which include RNeasy (Qiagen, Valencia, California) and StrataPrep (Stratagene, La Jolla, California). In an alternative embodiment, which is preferred for *S. cerevisiae*, RNA is extracted from cells using phenol and chloroform, as described in Ausubel *et al.*, eds., 1989, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Vol. III, Green Publishing Associates, Inc., John Wiley & Sons, Inc., New York, at pp. 13.12.1-13.12.5). Poly(A)⁺ RNA can be selected, *e.g.*, by selection with oligo-dT cellulose or, alternatively, by oligo-dT primed reverse transcription of total cellular RNA. In one embodiment, RNA can be fragmented by methods known in the art, *e.g.*, by incubation with ZnCl₂, to generate fragments of RNA. In another embodiment, the polynucleotide molecules analyzed by the invention comprise cDNA, or PCR products of amplified RNA or cDNA.

[00188] In one embodiment, total RNA, mRNA, or nucleic acids derived therefrom, is isolated from a sample taken from a person afflicted with breast cancer. Target polynucleotide molecules that are poorly expressed in particular cells may be enriched using normalization techniques (Bonaldo *et al.*, 1996, *Genome Res.* 6:791-806).

[00189] As described above, the target polynucleotides are detectably labeled at one or more nucleotides. Any method known in the art may be used to detectably label the target polynucleotides. Preferably, this labeling incorporates the label uniformly along the length of the RNA, and more preferably, the labeling is carried out at a high degree of efficiency. One embodiment for this labeling uses oligo-dT primed reverse transcription to incorporate the label; however, conventional methods of this method are biased toward generating 3' end fragments. Thus, in a preferred embodiment, random primers (*e.g.*, 9-mers) are used in reverse transcription to uniformly incorporate labeled nucleotides over the full length of the

target polynucleotides. Alternatively, random primers may be used in conjunction with PCR methods or T7 promoter-based *in vitro* transcription methods in order to amplify the target polynucleotides.

[00190] In a preferred embodiment, the detectable label is a luminescent label. For example, fluorescent labels, bioluminescent labels, chemiluminescent labels, and colorimetric labels may be used in the present invention. In a highly preferred embodiment, the label is a fluorescent label, such as a fluorescein, a phosphor, a rhodamine, or a polymethine dye derivative. Examples of commercially available fluorescent labels include, for example, fluorescent phosphoramidites such as FluorePrime (Amersham Pharmacia, Piscataway, N.J.), Fluoredate (Millipore, Bedford, Mass.), FAM (ABI, Foster City, Calif.), and Cy3 or Cy5 (Amersham Pharmacia, Piscataway, N.J.). In another embodiment, the detectable label is a radiolabeled nucleotide.

[00191] In a further preferred embodiment, target polynucleotide molecules from a patient sample are labeled differentially from target polynucleotide molecules of a standard. The standard can comprise target polynucleotide molecules from normal individuals (*i.e.*, those not afflicted with breast cancer). In a highly preferred embodiment, the standard comprises target polynucleotide molecules pooled from samples from normal individuals or tumor samples from individuals having sporadic-type breast tumors. In another embodiment, the target polynucleotide molecules are derived from the same individual, but are taken at different time points, and thus indicate the efficacy of a treatment by a change in expression of the markers, or lack thereof, during and after the course of treatment (*i.e.*, chemotherapy, radiation therapy or cryotherapy), wherein a change in the expression of the markers from a poor prognosis pattern to a good prognosis pattern indicates that the treatment is efficacious. In this embodiment, different timepoints are differentially labeled.

5.5.2.5 HYBRIDIZATION TO MICROARRAYS

[00192] Nucleic acid hybridization and wash conditions are chosen so that the target polynucleotide molecules specifically bind or specifically hybridize to the complementary polynucleotide sequences of the array, preferably to a specific array site, wherein its complementary DNA is located.

[00193] Arrays containing double-stranded probe DNA situated thereon are preferably subjected to denaturing conditions to render the DNA single-stranded prior to contacting with the target polynucleotide molecules. Arrays containing single-stranded probe DNA (*e.g.*, synthetic oligodeoxyribonucleic acids) may need to be denatured prior to contacting with the

target polynucleotide molecules, *e.g.*, to remove hairpins or dimers which form due to self complementary sequences.

[00194] Optimal hybridization conditions will depend on the length (*e.g.*, oligomer versus polynucleotide greater than 200 bases) and type (*e.g.*, RNA, or DNA) of probe and target nucleic acids. One of skill in the art will appreciate that as the oligonucleotides become shorter, it may become necessary to adjust their length to achieve a relatively uniform melting temperature for satisfactory hybridization results. General parameters for specific (*i.e.*, stringent) hybridization conditions for nucleic acids are described in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989), and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994). Typical hybridization conditions for the cDNA microarrays of Schena *et al.* are hybridization in 5 X SSC plus 0.2% SDS at 65°C for four hours, followed by washes at 25°C in low stringency wash buffer (1 X SSC plus 0.2% SDS), followed by 10 minutes at 25°C in higher stringency wash buffer (0.1 X SSC plus 0.2% SDS) (Schena *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:10614 (1993)). Useful hybridization conditions are also provided in, *e.g.*, Tijessen, 1993, HYBRIDIZATION WITH NUCLEIC ACID PROBES, Elsevier Science Publishers B.V.; and Kricka, 1992, NONISOTOPIC DNA PROBE TECHNIQUES, Academic Press, San Diego, CA.

[00195] Particularly preferred hybridization conditions include hybridization at a temperature at or near the mean melting temperature of the probes (*e.g.*, within 51°C, more preferably within 21°C) in 1 M NaCl, 50 mM MES buffer (pH 6.5), 0.5% sodium sarcosine and 30% formamide.

5.5.2.6 SIGNAL DETECTION AND DATA ANALYSIS

[00196] When fluorescently labeled probes are used, the fluorescence emissions at each site of a microarray may be, preferably, detected by scanning confocal laser microscopy. In one embodiment, a separate scan, using the appropriate excitation line, is carried out for each of the two fluorophores used. Alternatively, a laser may be used that allows simultaneous specimen illumination at wavelengths specific to the two fluorophores and emissions from the two fluorophores can be analyzed simultaneously (*see* Shalon *et al.*, 1996, "A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization," *Genome Research* 6:639-645, which is incorporated by reference in its entirety for all purposes). In a preferred embodiment, the arrays are scanned with a laser

fluorescent scanner with a computer controlled X-Y stage and a microscope objective. Sequential excitation of the two fluorophores is achieved with a multi-line, mixed gas laser and the emitted light is split by wavelength and detected with two photomultiplier tubes. Fluorescence laser scanning devices are described in Schena *et al.*, *Genome Res.* 6:639-645 (1996), and in other references cited herein. Alternatively, the fiber-optic bundle described by Ferguson *et al.*, *Nature Biotech.* 14:1681-1684 (1996), may be used to monitor mRNA abundance levels at a large number of sites simultaneously.

[00197] Signals are recorded and, in a preferred embodiment, analyzed by computer, *e.g.*, using a 12 or 16 bit analog to digital board. In one embodiment the scanned image is despeckled using a graphics program (*e.g.*, Hijaak Graphics Suite) and then analyzed using an image gridding program that creates a spreadsheet of the average hybridization at each wavelength at each site. If necessary, an experimentally determined correction for “cross talk” (or overlap) between the channels for the two fluors may be made. For any particular hybridization site on the transcript array, a ratio of the emission of the two fluorophores can be calculated. The ratio is independent of the absolute expression level of the cognate gene, but is useful for genes whose expression is significantly modulated in association with the different breast cancer-related condition.

5.6 THERAPEUTIC REGIMENS SPECIFIC TO PATIENT SUBSETS

[00198] The benefit of identifying subsets of individuals that have a common condition, followed by identification of sets of genes informative for those particular subsets of individuals, is that such subdivision and identification tends to more accurately identify the subset of genes responsible for, or most closely associated with, a particular form of the condition. For example, breast cancer is a complex condition brought about by several different molecular mechanisms. ER⁺ individuals, particularly ER⁺, ER/AGE high individuals, show an increased level of expression of cell cycle-control genes, and the expression of these genes is highly informative for prognosis in this patient subset (*see Examples*). In ER⁻ individuals, however, the expression of these genes is not informative for prognosis.

[00199] The set of informative markers, therefore, can be used to assign a particular course of therapy to an individual, *e.g.*, an individual having breast cancer, depending upon the condition subset into which the individual is classified. In one embodiment, therefore, the invention provides a method of assigning a course of therapy to an individual having a condition, said method comprising classifying the individual into one of a plurality of subsets

of a condition, wherein a plurality of informative genes has been identified for at least one of said subsets; and assigning a course of therapy known or suspected to be effective for treating the subset of the condition associated with those genes. In a specific embodiment, said condition is breast cancer, said patient subset is ER+, ER/AGE high status, and said course of therapy comprises the administration of one or more compounds known or suspected to be effective at arresting the cell cycle. In a more specific embodiment, said one or more compounds comprises taxol or a vinca alkaloid.

[00200] Of course, any course of therapy selected or assigned on the basis of the above phenotypes and gene expression may be supplemented by other treatments or courses of therapy relevant to or known or suspected to be effective in the treatment of the condition. For example, the treatment of breast cancer may additionally comprise surgery, either tissue-preserving or radical, radiation treatment, chemotherapy other than that suggested by gene expression analysis, or any other therapy or treatment known or suspected to be effective.

5.7 CLINICAL TRIALS AND EPIDEMIOLOGICAL STUDIES

[00201] The method of the present invention may also be used to assign individuals to categories within a clinical trial, epidemiological study or the like. For example, individuals may be distinguished according to a characteristic of a condition, such as the presence or absence of specific proteins (*e.g.*, estrogen receptor) or tissue structures (*e.g.*, lymph nodes), and with prognosis, and the results of the trial correlated with prognosis. In a specific example, the condition is breast cancer, the characteristic is the presence of the estrogen receptor, and the outcome is prognosis is the expected reoccurrence or non-reoccurrence of metastases within a given period, for example, five years, after initial diagnosis. In another specific example, the condition is obesity, the characteristics are 24-hour energy expenditure, and the prognosis is the expected occurrence of heart disease or diabetes. In another specific example, the condition is a neurodegenerative disease, the characteristic is exposure to a particular range of concentration of an environmental toxin, and the prognosis is expected occurrence or degree of loss of motor function. In each case, the characteristics and expected outcome are used to assign the individual to a category within a clinical trial or epidemiological study.

[00202] Thus, the invention provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising classifying the individual into one of a plurality of condition categories differentiated by at least one genotypic or phenotypic characteristic of the condition; determining the level of expression, in a sample derived from

said individual, of a plurality of genes informative for said condition category; determining whether said level of expression of said plurality of genes indicates that the individual has a good prognosis or a poor prognosis; and assigning the individual to a category in a clinical trial on the basis of prognosis.

[00203] In a specific embodiment, the invention provides a method of assigning an individual to a category in a breast cancer clinical trial, said method comprising: (a) classifying said individual as ER⁻, *BRCA1*, ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining for said individual the level of expression of at least two genes for which markers are listed in Table 1 if said individual is classified as ER⁻, *BRCA1*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻; (c) determining whether said individual has a pattern of expression of said at least two genes that correlates with a good prognosis or a poor prognosis; and (d) assigning said individual to at least one category in a clinical trial if said individual has a good prognosis, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis. In a more specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a). In another more specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of any other clinical, phenotypic or genotypic characteristic of breast cancer. In another more specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis of breast cancer, and determining from the expression of said second plurality of genes, in addition to said first plurality of genes, whether said individual has a good prognosis or a poor prognosis.

5.8 KITS

[00204] The present invention further provides for kits comprising the marker sets described above. The components of the kits of the present invention are preferably contained in sealed containers. In a preferred embodiment, the kit comprises a microarray ready for hybridization to target polynucleotide molecules. In specific embodiments, the kit may comprise any of the microarrays described in detail in Section 5.5.2. Where proteins are the target molecules, the kit preferably comprises a plurality of antibodies for binding to specific

condition-related proteins, and means for identifying such binding (*e.g.*, means for performing a sandwich assay, ELISA, RIA, *etc.*). Such antibodies may be provided, for example, individually or as part of an antibody array. The kit may additionally comprise software for the data analyses described above, as described in detail in Section 5.9. The kit preferably contains one or more control samples. Such a control sample may be an artificial population of marker-related or marker-derived polynucleotides suitable for hybridization to a microarray, wherein the markers are related to or relevant to the condition of interest (for example, breast cancer). The control may also, or alternatively, be a set of expression values stored on a computer disk or other storage medium.

[00205] The kits of the invention may be primarily diagnostic in nature; that is, they may assist a physician or researcher in determining a characteristic, for example, the prognosis, of a condition of interest, the likely response to a therapeutic regimen, the likely outcome of exposure to an environmental condition, such as toxin exposure, *etc.* The kits of the invention may also be used to classify individuals, for example, to place individuals into different groups in a clinical trial. The use of each kit is determined by the markers, microarrays, controls, *etc.* included.

[00206] COMPUTER-FACILITATED ANALYSIS The analytic methods described in the previous sections can be implemented by use of the following computer systems and according to the following programs and methods. A computer system comprises internal components linked to external components. The internal components of a typical computer system include a processor element interconnected with a main memory. For example, the computer system can be based on an Intel 8086-, 80386-, 80486-, Pentium™, or Pentium™-based processor with preferably 32 MB or more of main memory. The computer system may also be a Macintosh or a Macintosh-based system, but may also be a minicomputer or mainframe.

[00207] The external components preferably include mass storage. This mass storage can be one or more hard disks (which are typically packaged together with the processor and memory). Such hard disks are preferably of 1 GB or greater storage capacity. Other external components include a user interface device, which can be a monitor, together with an inputting device, which can be a “mouse”, or other graphic input devices, and/or a keyboard. A printing device can also be attached to the computer.

[00208] Typically, a computer system is also linked to network link, which can be part of an Ethernet link to other local computer systems, remote computer systems, or wide area

communication networks, such as the Internet. This network link allows the computer system to share data and processing tasks with other computer systems.

[00209] Loaded into memory during operation of this system are several software components, which are both standard in the art and special to the instant invention. These software components collectively cause the computer system to function according to the methods of this invention. These software components are typically stored on the mass storage device. A software component comprises the operating system, which is responsible for managing computer system and its network interconnections. This operating system can be, for example, of the Microsoft Windows[®] family, such as Windows 3.1, Windows 95, Windows 98, Windows 2000, or Windows NT, or may be of the Macintosh OS family, or may be UNIX, a UNIX derivative such as LINUX, or an operating system specific to a minicomputer or mainframe. The software component represents common languages and functions conveniently present on this system to assist programs implementing the methods specific to this invention. Many high or low level computer languages can be used to program the analytic methods of this invention. Instructions can be interpreted during run-time or compiled. Preferred languages include C/C++, FORTRAN and JAVA. Most preferably, the methods of this invention are programmed in mathematical software packages that allow symbolic entry of equations and high-level specification of processing, including some or all of the algorithms to be used, thereby freeing a user of the need to procedurally program individual equations or algorithms. Such packages include Matlab from Mathworks (Natick, MA), Mathematica[®] from Wolfram Research (Champaign, IL), or S-Plus[®] from Math Soft (Cambridge, MA). Specifically, the software component includes the analytic methods of the invention as programmed in a procedural language or symbolic package.

[00210] The software to be included with the kit comprises the data analysis methods of the invention as disclosed herein. In particular, the software may include mathematical routines for marker discovery, including the calculation of similarity values between clinical categories (*e.g.*, prognosis) and marker expression. The software may also include mathematical routines for calculating the similarity between sample marker expression and control marker expression, using array-generated fluorescence data, to determine the clinical classification of a sample.

[00211] Additionally, the software may also include mathematical routines for determining the prognostic outcome, and recommended therapeutic regimen, for an individual with a condition of interest. In the specific example of breast cancer, the mathematical routines

would determine the prognostic outcome and recommended therapeutic regimen for an individual having breast cancer. Such breast cancer-specific software would include instructions for the computer system's processor to receive data structures that include the level of expression of five or more of the marker genes listed in any of Tables 1-5 in a breast cancer tumor sample obtained from the breast cancer patient; the mean level of expression of the same genes in a control or template; and the breast cancer patient's clinical information, including age, lymph node status and ER status. The software may additionally include mathematical routines for transforming the hybridization data and for calculating the similarity between the expression levels for the marker genes in the patient's breast cancer tumor sample and a control or template. In a specific embodiment, the software includes mathematical routines for calculating a similarity metric, such as a coefficient of correlation, representing the similarity between the expression levels for the marker genes in the patient's breast cancer tumor sample and the control or template, and expressing the similarity as that similarity metric.

[00212] The software preferably would include decisional routines that integrate the patient's clinical and marker gene expression data, and recommend a course of therapy. In one embodiment, for example, the software causes the processor unit to receive expression data for prognosis-related genes in the patient's tumor sample, calculate a metric of similarity of these expression values to the values for the same genes in a template or control, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to the prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen. In a specific example, the software additionally causes the processor unit to receive data structures comprising clinical information about the breast cancer patient. In a more specific example, such clinical information includes the patient's age, estrogen receptor status, and lymph node status.

[00213] The software preferably causes the processor unit to receive data structures comprising relevant phenotypic and/or genotypic characteristics of the particular condition of interest, and/or of an individual having that condition, and classifies the individual into a condition subset according to those characteristics. The software then causes the processor to receive values for subset-specific markers, to calculate a metric of similarity of the values associated with those markers (*e.g.*, level, abundance, activity, *etc.*) from the individual to a control, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to a prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen. In the

specific example of breast cancer and a breast cancer patient, the software, in one embodiment, causes the processor unit to receive data structures comprising the patient's age, estrogen receptor status, and lymph node status, and on the basis of this data, to classify the patient into one of the following patient subsets: ER⁻, sporadic; ER⁻, *BRCAl*; ER⁺, AR/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻. The software then causes the processor to receive expression values for subset-specific prognosis-informative gene expression in the patient's tumor sample, calculate a metric of similarity of these expression values to the values for the same genes in a patient subset-specific template or control, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to the prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen.

[00214] Where the control is an expression template comprising expression values for marker genes within a group of patients, *e.g.*, breast cancer patients, the control can comprise either hybridization data obtained at the same time (*i.e.*, in the same hybridization experiment) as the patient's individual hybridization data, or can be a set of hybridization or marker expression values stored on a computer, or on computer-readable media. If the latter is used, new patient hybridization data for the selected marker genes, obtained from initial or follow-up tumor samples, or suspected tumor samples, can be compared to the stored values for the same genes without the need for additional control hybridizations. However, the software may additionally comprise routines for updating the control data set, *e.g.*, to add information from additional breast cancer patients or to remove existing members of the control data set, and, consequently, for recalculating the average expression level values that comprise the template. In another specific embodiment, said control comprises a set of single-channel mean hybridization intensity values for each of said at least five of said genes, stored on a computer-readable medium.

[00215] Clinical data relating to a breast cancer patient, or a patient having another type of condition, and used by the computer program products of the invention, can be contained in a database of clinical data in which information on each patient is maintained in a separate record, which record may contain any information relevant to the patient, the patient's medical history, treatment, prognosis, or participation in a clinical trial or study, including expression profile data generated as part of an initial diagnosis or for tracking the progress of the condition, for example, breast cancer, during treatment.

[00216] Thus, one embodiment of the invention provides a computer program product for classifying a breast cancer patient according to prognosis, the computer program product for

use in conjunction with a computer having a memory and a processor, the computer program product comprising a computer readable storage medium having a computer program mechanism encoded thereon, wherein said computer program product can be loaded into the one or more memory units of a computer and causes the one or more processor units of the computer to execute the steps of (a) receiving a first data structure comprising said breast cancer patient's age, ER status, LN status and tumor type; (b) classifying said patient as ER⁻, sporadic; ER⁻, *BRCAl*; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (c) receiving a first data structure comprising the level of expression of at least two genes in a cell sample taken from said breast cancer patient wherein markers for said at least two genes are listed in Table 1 if said patient is classified as ER⁻, sporadic; Table 2 if said patient is classified as ER⁻, sporadic; Table 3 if said patient is classified as ER⁺, ER/AGE high; Table 4 if said patient is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said patient is classified as ER⁺, ER/AGE high, LN⁻; (d) determining the similarity of the level of expression of said at least two genes to control levels of expression of said at least two genes to obtain a patient similarity value; (e) comparing said patient similarity value to selected first and second threshold values of similarity of said level of expression of said genes to said control levels of expression to obtain first and second similarity threshold values, respectively, wherein said second similarity threshold indicates greater similarity to said control levels of expression than does said first similarity threshold; and (f) classifying said breast cancer patient as having a first prognosis if said patient similarity value exceeds said first and said second threshold similarity values, a second prognosis if said patient similarity value exceeds said first threshold similarity value but does not exceed said second threshold similarity value, and a third prognosis if said patient similarity value does not exceed said first threshold similarity value or said second threshold similarity value. In a specific embodiment of said computer program product, said first threshold value of similarity and said second threshold value of similarity are values stored in said computer. In another more specific embodiment, said first prognosis is a "very good prognosis," said second prognosis is an "intermediate prognosis," and said third prognosis is a "poor prognosis," and wherein said computer program mechanism may be loaded into the memory and further cause said one or more processor units of said computer to execute the step of assigning said breast cancer patient a therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and expression profile. In another specific embodiment, said computer program mechanism may

be loaded into the memory and further cause said one or more processor units of the computer to execute the steps of receiving a data structure comprising clinical data specific to said breast cancer patient. In a more specific embodiment, said single-channel hybridization intensity values are log transformed. The computer implementation of the method, however, may use any desired transformation method. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating the difference between the level of expression of each of said genes in said cell sample taken from said breast cancer patient and the level of expression of the same genes in said control. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating the mean log level of expression of each of said genes in said control to obtain a control mean log expression level for each gene, calculating the log expression level for each of said genes in a breast cancer sample from said breast cancer patient to obtain a patient log expression level, and calculating the difference between the patient log expression level and the control mean log expression for each of said genes. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating similarity between the level of expression of each of said genes in said cell sample taken from said breast cancer patient and the level of expression of the same genes in said control, wherein said similarity is expressed as a similarity value. In more specific embodiment, said similarity value is a correlation coefficient. The similarity value may, however, be expressed as any art-known similarity metric.

[00217] Of course, the above breast cancer-specific examples are not limiting; analogous computer systems, software, and data analysis methods may be utilized for any condition of interest. For example, analogous software may be used to determine the prognosis of any other type of cancer, or of any other non-cancer diseases or conditions, using markers, expression level data and controls specific for that cancer, non-cancer disease or condition.

[00218] In an exemplary implementation, to practice the methods of the present invention, a user first loads experimental data into the computer system. These data can be directly entered by the user from a monitor, keyboard, or from other computer systems linked by a network connection, or on removable storage media such as a CD-ROM, floppy disk (not illustrated), tape drive (not illustrated), ZIP[®] drive (not illustrated) or through the network. Next the user causes execution of expression profile analysis software which performs the methods of the present invention.

[00219] In another exemplary implementation, a user first loads experimental data and/or databases into the computer system. This data is loaded into the memory from the storage media or from a remote computer, preferably from a dynamic geneset database system, through the network. Next the user causes execution of software that performs the steps of the present invention.

[00220] Additionally, because the data obtained and analyzed in the software and computer system products of the invention may be confidential, the software and/or computer system preferably comprises access controls or access control routines, such as password protection and preferably, particularly if information is to be transmitted between computers, for example, over the Internet, encryption of the data by a suitable encryption algorithm (*e.g.*, PGP).

[00221] Alternative computer systems and software for implementing the analytic methods of this invention will be apparent to one of skill in the art and are intended to be comprehended within the accompanying claims. In particular, the accompanying claims are intended to include the alternative program structures for implementing the methods of this invention that will be readily apparent to one of skill in the art.

6. EXAMPLE: IDENTIFICATION OF PHENOTYPIC SUBSETS AND INFORMATIVE GENESETS FOR EACH

[00222] Materials and Methods

Tumor Samples:

[00223] 311 cohort samples were collected from breast cancer patients. Selection criteria for sporadic patients (*i.e.*, those not identified as having a *BRCA1*-type tumor; $n = 291$) included: primary invasive breast carcinoma less than 5 cm (T1 or T2); no axillary metastases (N0); age at diagnosis of less than 55 years; calendar year of diagnosis 1983-1996; and no previous malignancies. All patients were treated by modified radical mastectomy or breast-conserving treatment. *See van't Veer et al., Nature* 415:530 (2002). Selection criteria for hereditary (*i.e.*, *BRCA1*-type; $n = 20$) tumors included: carriers of germline mutation in *BRCA1* or *BRCA2*, and primary invasive breast carcinoma. *van't Veer, supra*. Additionally, for development of a classifier for the *BRCA1* group, 14 *BRCA1* samples previously identified (*see van't Veer, supra*) were added to the 20 *BRCA1* type samples to increase sample size. Those 14 samples also satisfy the conditions that they are ER negative and age less than 55 years old.

[00224] Data analysis:

[00225] Sample sub-grouping: As shown in FIG. 1, tumor samples were first divided into ER⁺ and ER⁻ branches since this is the dominant gene expression pattern. In the ER⁻ branch, the samples were further divided into “BRCA1 mutation like” and “Sporadic like” categories using the expression templates and 100 genes previously identified as optimal for determining *BRCA1* status. See van’t Veer *et al.*, *Nature* 415:530 (2002). In the ER⁺ category, samples were divided by ER vs. age distribution (see below) into two groups, “ER/AGE low” and “ER/AGE high.” Within the “ER/AGE low” group, samples were further divided according to the lymph node status into two sub-groups: lymph node negative (0 lymph nodes; LN⁻) and positive (> 0 lymph nodes; LN⁺) group.

[00226] The result of these divisions was five distinctive sub-groups: “ER⁻, sporadic” ($n = 52$), “ER⁻, BRCA1” ($n = 34$), “ER⁺, ER/AGE high” ($n = 83$), “ER⁺, ER/AGE low, LN⁻” ($n = 81$), and “ER⁺, ER/AGE low, LN⁺” ($n = 75$). A few samples with a specific ER vs. age distribution in “ER⁺, ER/AGE low, LN⁺” group were further excluded to develop a classifier, see below for details.

[00227] Estrogen receptor level: Estrogen receptor gene expression level was measured by a 60mer oligo-nucleotide on a microarray. Since every individual sample was compared to a pool of all samples, the ratio to pool was used to measure the relative level. A threshold of -0.65 on $\log_{10}(\text{ratio})$ was used to separate the ER⁺ group from ER⁻ group. See van’t Veer *et al.*, *Nature* 415:530 (2002).

[00228] Grouping by ER vs. age distribution: Samples were not uniformly distributed in ER vs. age space among the ER⁺ samples (FIG. 2). First, it appeared that the ER level increases with age, as there were few samples from young individuals having a high ER expression level. For example, in the 35 to 40 years age group, samples having a $\log(\text{ratio})$ of ER > 0.2 are relatively few as compared to the 40 to 45 age group. In the set of samples used, the $40 < \text{age} \leq 45$ group contains 30 samples having $\log(\text{ratio})$ ER values between -0.2 to 0.2 , and 28 samples having values greater than 0.2 , whereas the $35 < \text{age} \leq 40$ group includes 24 samples with values between -0.2 to 0.2 , but only 6 samples with values of greater than 0.2 (Fisher’s exact test P-value: 1%). The increasing ER level with age may simply due to the fact that estrogen levels decrease with age, and the estrogen receptor level rises in compensation.

[00229] There also appeared to be at least two groups of patients, as indicated by the solid line separating the two in FIG. 2A. A bimodality test of the separation indicated by the solid line yielded P-value $< 10^{-4}$. Each of these two groups has its own trend between the ER level and age. The solid line can be approximated by $\text{ER} = 0.1(\text{age} - 42.5)$. Patients having values

above the solid line are referred to as the “ER/AGE high” group, and the patients below the line as the “ER/AGE low” group.

[00230] Prognosis in each group:

[00231] Feature selection and performance evaluation: For the prognosis in each group, non-informative genes were filtered in each group of patients. Specifically, only genes with $|\log_{10}(\text{ratio})| > \log_{10}(2)$ and P-value (for $\log(\text{ratio}) \neq 0$) < 0.01 in more than 3 experiments were kept. This step removed all genes that never had any significant change across all samples. The second step used a leave-one-out cross validation (LOOCV) procedure to optimize the number of reporter genes (features) in the classifier and to estimate the performance of the classifier in each group. The feature selection was included inside the loop of each LOOCV process. The final “optimal” reporter genes were selected using all of the “training samples” as the result of “re-substitution” because one classifier was needed for each group.

[00232] Selection of training samples: Only the samples from patients who had metastases within 5 years of initial diagnosis (3 years for “ER⁻, sporadic” samples; *i.e.*, the “poor outcome” group), or who were metastases-free with more than 5 years of follow-up time (*i.e.*, the “good outcome” group), were used as the training set. Because the average expression levels for informative genes among patients who were metastasis-free, or who had early metastases, were used as expression templates for prediction, the training samples for the ER⁺ samples were further limited to those samples that could also be correctly classified by the first round of LOOCV process. For the “ER⁻, sporadic” samples, no such iteration was done because no improvement was observed. For the “ER⁻, BRCA1” samples, an iteration was done, but the training samples in the second iteration were limited to the correctly predicted good outcome samples from the first round of LOOCV, and all the poor outcome samples with metastases time less than 5 years. Further limitation of the poor outcome samples was not performed because of the small number of poor samples and the absence of improvement by such limitation. In the first round of LOOCV, except for the “ER⁻, sporadic” group, the number of features was fixed at 50 genes. A patient was predicted to have a favorable outcome, that is, no metastases within five years of initial diagnosis, if the expression of the reporter genes in a sample from the individual was more similar to the “average good profile” than the “average poor profile”, and a poor outcome, that is, a metastasis within five years, if the expression of the reporter genes in the sample was more similar to the “average poor profile” than the “average good profile”.

[00233] The justification for such an iteration operation is threefold. First, biologically, there are always a few individuals with specific reasons (different from the vast majority) to stay metastases free or to develop metastases. Second, statistically, most groups of patients include outliers that don't follow the distribution of the majority of samples. Third, methodologically, the iteration operation is very similar to the idea of "boosting", but instead of increasing the weights of the samples predicted wrong, emphasis is placed on the well behaved samples for selecting features and training the classifier. Since this process was used to select "training samples", and the performance was evaluated using the LOOCV (including the feature selection) after the training sample being fixed, there is no issue of over-fitting involved in our procedures. This method of iteration is thus more likely to reveal the dominant mode to metastases within each group.

[00234] Error rate and odds ratio, threshold in the final LOOCV: Unless otherwise stated, the error rate was the average error rate from two populations: (1) the number of poor outcome samples misclassified as good outcome samples, divided by the total number of poor outcome samples; and (2) the total number of good outcome samples misclassified as poor outcome samples, divided by the total number of good samples. Two odds ratios were reported for a given threshold: (1) the overall odds ratio and (2) the 5 year odds ratio. The 5 year odds ratio was calculated from samples from individuals that were metastases free for more than five years, and who experienced metastasis within 5 years. The threshold was applied to **cor1** — **cor2**, where "cor1" stands for the correlation to the "average good profile" in the training set, and "cor2" stands for the correlation to the "average poor profile" in the training set.

[00235] The threshold in the final round of LOOCV was defined using the following steps: (1) For each of the N sample i left out for training, features based on the training set were selected, (2) given a feature set, an incomplete LOOCV with $N-1$ samples was performed (only the "average poor profile" and "average good profile" is varied depending on whether the left out sample is in the training set or not), (3) the threshold based on the minimum error rate from $N-1$ samples was determined, and that threshold was assigned to sample i in step (1), (4) the median threshold from all N samples was taken, and designated the final threshold. FIGS. 3-7 present detailed information about classifiers for the 5 groups: "ER⁻, sporadic", "ER⁻, BRCA1", "ER⁺, ER/age high", "ER⁺, ER/age low, LN⁻", "ER⁺, ER/age low, LN⁺". Tables 1-5 (*see* Section 5.3) list the final optimal reporter genes for each of the 5 classifiers for each of the five patient subsets. Table 6, below, summarizes the performance of each of the five classifiers together with thresholds used in each classifier.

[00236] Table 6. Performance of classifiers for each patient subset.

Classifier	Optimal # of Genes	(C1-C2) Threshold	Metastasis Free	# of Samples	TP	FP	FN	TN	Odds Ratio	95% C.I.
ER+, ER/AGE high	50	1.22	Overall	83	31	14	5	33	14.61	4.71-45.36
			5 year	71	24	11	3	33	24.00	6.03-95.46
ER+, ER/AGE low, LN-	65	0.38	Overall	81	14	6	6	55	21.39	5.98-76.52
			5 year	73	11	4	5	53	29.15	6.73-126.33
ER+, ER/AGE low, LN+	50	-0.12	Overall	56	7	4	6	39	11.38	2.54-50.94
			5 year	48	5	4	3	36	15.00	2.57-87.64
ER-, sporadic	20	-0.01	Overall	52	18	7	7	29	7.35	2.16-25.04
			5 year	45	16	5	6	18	9.60	2.45-37.58
ER-, BRCA1	10	-0.37	Overall	34	6	3	3	22	14.67	2.34-92.11
			5 year	22	6	1	3	12	24.00	2.04-282.68

[00237] TP: True positive

[00238] FP: False positive

[00239] FN: False negative

[00240] TN: True negative

[00241] Classification method: All classifiers described herein, feature selection and optimization were included inside the LOOCV loop. Classifier performance was based on the LOOCV results. The profile based on the selected features from each patient was compared to the “average good profile” and “average poor profile” (by correlation) to determine its predicted outcome.

[00242] Correlation calculation: The correlation between each gene’s expression log(ratio) and the endpoint data (final outcome) was calculated using the Pearson’s correlation coefficient. The correlation between each patient’s profile and the “average good profile” and “average poor profile” was the cosine product (no mean subtraction).

[00243] Results:

[00244] The comprehensive prognosis strategy was employed on microarray expression profiles of 311 patients diagnosed before age 55 that were all part of previous studies establishing and validating a 70-gene prognosis profile. See van ’t Veer *et al.*, *Nature* 415:530 (2002); van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002). In addition, 14 known *BRCA1* samples from the *Nature* study were included in defining the prognosis

classifier for the *BRCA1* group. The overview of the stratifications is shown in FIG. 1. In each of the patient subsets, prognosis classifiers were developed and performance was evaluated by leave-one-out cross-validation. The biological make up of each of the classifiers was also examined.

[00245] During the process to decide whether a particular clinical parameter should be used for the next stratification, our objectives were twofold: (1) identification of homogeneous prognosis patterns; and/or (2) improved prognosis in the subsets. There is a subtle balance between these two objectives because smaller groups will likely lead to uniform patterns within the group but have increasingly limited predictive power. With the exception of the *BRCA1* subset, each group in our stratification contained 50 or more samples.

[00246] The first layer of stratification was based on the estrogen receptor level. It was previously observed that estrogen receptor expression has a dominant effect on overall gene expression in breast cancer as seen in hierarchical clustering. van 't Veer *et al.*, *Nature* 415:530 (2002); Perou *et al.*, *Nature* 406:747 (2000); Gruvberger *et al.*, *Cancer Res.* 61:5979 (2001). In previous analysis up to 2500 genes were significantly correlated with ER expression levels in tumor. See, van 't Veer *et al.*, *Nature* 415:530 (2002). According to the threshold defined previously (van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002)), samples were first divided into two groups according to the estrogen receptor level as measured by the oligo probe (accession number: NM_000125) on the array; samples with $\log(\text{ratio}) > -0.65$ belong to the ER⁺ group, and the rest belong to ER⁻ group). This resulted in 239 samples in the ER⁺ group and 72 samples in the ER⁻ group.

[00247] In the ER⁺ branch it was observed that when displaying ER expression level as a function of age, at least two subgroups appeared to exist. (In general, any bimodality in the clinical data is useful.) The tumors were stratified according this bimodality (*see* FIG. 2). The group of ER⁺ patients having a high ER/AGE ratio was designated the “ER/AGE high” group (83 samples), and the remaining group of patients was designated “ER/AGE low” group (156 samples).

[00248] Within the “ER/age high” group, a group of prognosis reporter genes that highly correlated with the outcome is identified (*see* Table 3). Moreover, the expression of these genes appeared to be very homogeneous, as indicated by high similarity in expression among those genes. *See* FIG 2A. Leave-one-out cross validation including reporter selection yielded an odds ratio of 14.6 (95%CI: 4.7-45.4) and 5 year odds ratio of 24.0 (95%CI: 6.0-95.5). Examination of those reporter genes reveals they are mostly the cell cycle genes which are highly expressed in the poor outcome tumors. It is worth noting that even though this

group includes LN+ and LN- individuals, and mixed treatment, the incidence of distant metastases is predicted by a biologically uniform set of genes, possibly indicating that proliferation is the prime driving force for disease progression. Also even though variation in these genes is observed in other tumor subgroups this is generally not correlated with outcome in those settings (see below).

[00249] In the “ER/age low” group, no predictive pattern was found in the whole group; thus, the samples were further stratified into LN- (81 samples, referred to as “ER/age low LN⁻”) and LN+ (75 samples, referred to as “ER/age low LN⁺”) group.

[00250] Within the “ER/age low LN⁻” group, a group of genes was identified that was uniformly co-regulated, and which correlated with the outcome. Leave-one-out cross-validation (including feature selection) yielded an odds ratio of 21.4 (95% CI: 6.0-76.5) and 5 year odds ratio of 29.2 (95% CI: 6.7-126.3). This group of genes is also enriched for individual biological functions (see below).

[00251] For the “ER/age low LN⁺” subset, an informative set of genes (*see* Table 4) was obtained after exclusion of several samples from older individuals having low ER levels. These samples are indicated in FIG. 2A as those lying below the dashed line (approximated as $ER < 0.1 * (age - 50)$). 56 samples remained after the exclusion. This sample set allowed the identification of a group of genes with a highly homogeneous pattern that is useful for prognosis (overall odds ratio: 11.4 (2.5-50.9), 5 year odds ratio: 15.0 (2.6-87.6)). This suggests again that ER vs. age is an important combination for stratifying breast cancer patients. The reporter genes involved in this classifier also correlated with the clinical measure of the degree of lymphocytic infiltration (data not shown). The prediction in this group was not as strong as other positive groups, which may indicate the primary tumor carries weaker information about the metastases for this group of patients, and the metastases may be started from or influenced by tumors already in lymph nodes.

[00252] In the ER⁻ branch, because a portion of the samples are “*BRCA1*-like,” it is natural to divide the samples into “*BRCA1*-like” and “sporadic like”. To perform the classification, the *BRCA1*/sporadic tumor type classifier described in Roberts *et al.*, “Diagnosis and Prognosis of Breast Cancer Patients,” International Publication No. WO 02/103320, which is hereby incorporated by reference in its entirety, to segregate the ER⁻ cohort samples. 52 out of the 72 ER⁻ samples were found to be “sporadic like” and 20 were found to be “*BRCA1*-like”. Interestingly, the “sporadic like” group was enriched for *erbb2* mutations (data not shown).

[00253] Within the “ER⁻, sporadic” group, no homogeneous prognosis pattern was identified; however, 20 genes were identified that are highly predictive of the tumor outcome (see Table 2). Leave-one-out cross-validation including feature selection yielded an odds ratio of 7.4 (95% CI 2.2-25.0) and 5 year odds ratio 9.6 (2.5 – 37.6). This result represents a significant improvement in prognosis compared to the previously-identified 70 gene prognosis classifier (see Roberts *et al.*, International Publication No. WO 02/103320; van 't Veer *et al.*, *Nature* 415:530 (2002)) which has no within-group prognostic power for the ER⁻ patient subset. The fact that 20 genes predict outcome and that there is no homogeneous (and apparent biological) pattern in this group probably indicates multiple mechanisms of metastasis in this group. Gene annotation indicates that genes included may be involved in invasion, energy metabolism and other functions.

[00254] For the “ER⁻, *BRCA1*-like” group, we added 14 *BRCA1* mutation carrier samples from a previous study were added to increase the number of samples. Those 14 extra samples also satisfied the following selection criteria: ER negative and age less than 55 years. The leave-one-out cross validation process identified 10 genes that are predictive of final outcomes. The overall odds ratio is 14.7 (95% CI: 2.3-92.1) and the 5 year odds ratio is 24.0 (95% CI: 2.0-282.7).

[00255] Because no homogeneous gene expression patterns were found in ER⁻ branch, the predictive power of those genes was further validated. One means of further validation was to review the different classifier gene sets for biological interpretations and to identify genes within each classifier that gave indications as to the origins of the tumors.

[00256] The “ER⁺, ER/AGE high” group yielded a classifier highly enriched for cell cycle genes with both G1/S and G2/M phases represented. In this group, over-expression of 46 of the 50 genes was associated with disease progression including all the known cell cycle genes. This is consistent with rapid growth being the determinant of metastatic potential. Four genes in this classifier were anti-correlated with outcome and cell cycle. One of these genes encodes follistatin, which binds to and inhibits activin and other members of the TGFβ family (Lin *et al.*, *Reproduction* 126:133 (2003)), the members of which have many functions, including growth stimulation. Tumor grade also accurately predicted metastatic potential in this group (overall odds ratio: 5.9, 95% CI: 2.0-18.0, 5 year odds ratio: 12.5, 95% CI: 2.6-59.3) and was also correlated with the expression level of these genes, which is consistent with rate of growth being the primary determinant of disease progression. This set of genes had a significantly lower correlation with outcome in the other patient subsets, even though coordinate and similarly variable expression was seen. For example, many tumors in

the “ER⁻, sporadic” group had high cell cycle and low FST expression, but the expression of these genes in these groups was minimally correlated with outcome, indicating that growth was not the primary determinant of outcome here (*see* FIGS. 8A and 8B).

[00257] The ER⁺, ER/AGE low, LN⁻ group yielded a classifier rich in both genes for glycolytic enzymes (12 of 56) and genes induced by hypoxia and/or angiogenesis (14 of 56) with 5 genes falling into both categories. These genes were positively correlated with poor outcome, implying that energy metabolism (glycolysis), angiogenesis and adaptation to hypoxia were critical pathways in this subgroup of tumors. None of these genes appeared in the classifiers for the other patient subsets, and there was a much reduced predictive value of these genes in the other tumors, even though coordinate and similarly variable expression was seen (*see* FIG. 8C and 8D).

[00258] The implication of the above analyses is that certain well known functions (growth, angiogenesis, energy metabolism) are important in certain tumor types and not in others, and therefore therapies that target these functions will be likely be similarly effective in some tumor subgroups and not in others. For example therapies that target cell cycle progression, such as taxol or the vinca alkaloids, may be optimally effective in the ER⁺, ER/AGE high group, where overexpression of cell cycle genes predominates in the classifier. In contrast, tumor subgroups in which variation in cell cycle expression is not correlated with outcome may be less sensitive to taxol or the vinca alkaloids.

[00259] The “comprehensive prognosis” approach significantly improved the prediction error rate when compared with 70 gene classifier (Table 7). To make the comparison fair, we listed two sets of results from the 70 gene classifier. The first results from the use of the same threshold applied to all the patient subsets (threshold previously optimized for false negative rate); the second one results from the use of a threshold optimized for each patient subset (optimized for average error rate). The comprehensive approach lowered the error rate by at least 6%.

Table 7. Average error rate for the patient subset approach compared with the previously-described 70 gene classifier.

Prognosis method	over all error rate	5 year error rate
70 gene, fix thresh	30.90%	25.70%
70 gene, opt thresh	28.60%	27.60%
Comprehensive	21.50%	19.30%

[00260] Fix thresh: use of a fixed threshold in the classifier as previously determined.

[00261] Opt threshold: use of a threshold optimized for each sub-group. For the “ER/Age low, LN+” subgroup, 56 samples used for developing the classifier were included here, resulted in 306 samples in total.

[00262]

7. REFERENCES CITED

[00263] All references cited herein are incorporated herein by reference in their entirety and for all purposes to the same extent as if each individual publication or patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes.

[00264] Many modifications and variations of the present invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be limited only by the terms of the appended claims along with the full scope of equivalents to which such claims are entitled.

WHAT IS CLAIMED IS:

1. A method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising:

(a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics of said condition into a plurality of first classes; and

(b) identifying within each of said first classes a first set of genes or markers informative for said condition

wherein said first set of genes or markers within each of said first classes is unique to said class relative to other first classes.

2. The method of claim 1, which further comprises additionally classifying into a plurality of second classes said samples or individuals in at least one of said first classes on the basis of a phenotypic or genotypic characteristic different than that used in said classifying step (a); and identifying within at least one of said second classes a second set of informative genes or markers, wherein said second set of informative genes or markers within each of said second classes is unique to said second class relative to other first and second classes.

3. A method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising:

(a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics into a plurality of first classes;

(b) classifying at least one of said first classes into a plurality of second classes on the basis of phenotypic or genotypic characteristic different than that used in said classifying step (a); and

(c) identifying within at least one of said first classes or said second classes a set of genes or markers informative for said condition,

wherein said second set of genes or markers is unique to said class relative to other first and second classes.

4. A method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising:

(a) selecting a first characteristic from said plurality of phenotypic or genotypic characteristics;

(b) identifying at least two first condition classes differentiable by said first characteristic;

(c) selecting a plurality of individuals classifiable into at least one of said first condition classes; and

(d) identifying in samples derived from each of said plurality of individuals a set of genes or markers informative for said condition within said at least one of said first condition classes.

5. A method of classifying an individual with a condition as having a good prognosis or a poor prognosis, comprising:

(a) classifying said individual into one of a plurality of patient classes, said patient classes being differentiated by one or more phenotypic, genotypic or clinical characteristics of said condition;

(b) determining the level of expression of a plurality of genes or their encoded proteins in a cell sample taken from the individual relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins informative for prognosis of the patient class into which said individual is classified; and

(c) classifying said individual as having a good prognosis or a poor prognosis on the basis of said level of expression.

6. The method of claim 5, wherein said condition is cancer, said good prognosis is the non-occurrence of metastases within five years of initial diagnosis, and said poor prognosis is the occurrence of metastases within five years of initial diagnosis.

7. The method of claim 5, wherein said control is the average level of expression of each of said plurality of genes or their encoded proteins across a plurality of samples derived from individuals identified as having a poor prognosis.

8. The method of claim 7, in which said classifying step (c) is carried out by a method comprising comparing the level of expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a poor prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in said control more strongly than would be expected by chance.

9. The method of claim 5, wherein said control is the average level of expression of each of said plurality of genes or their encoded proteins across a plurality of samples derived from individuals identified as having a good prognosis.

10. The method of claim 9, in which said classifying in step (c) is carried out by a method comprising comparing the level expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a good prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in said control more strongly than would be expected by chance.

11. The method of claim 5, wherein said plurality of patient classes comprises ER^{-} , *BRCAI* individuals; ER^{-} , sporadic individuals; ER^{+} , *ER/AGE* high individuals; ER^{+} , *ER/AGE* low, *LN*⁺ individuals; and ER^{+} , *ER/AGE* low, LN^{-} individuals.

12. A method of classifying a breast cancer patient as having a good prognosis or a poor prognosis comprising:

(a) classifying said breast cancer patient as ER^{-} , *BRCAI*; ER^{-} , sporadic; ER^{+} , *ER/AGE* high; ER^{+} , *ER/AGE* low, *LN*⁺; or ER^{+} , *ER/AGE* low, LN^{-} ;

(b) determining the level of expression of a first plurality of genes in a cell sample taken from said breast cancer patient relative to a control, said first plurality of genes comprising two of the genes corresponding to the markers in Table 1 if said breast cancer patient is classified as ER^{-} , *BRCAI*; in Table 2 if said breast cancer

patient is classified as ER⁻ sporadic; in Table 3 if said breast cancer patient is classified as ER⁺, ER/AGE high; in Table 4 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁺; or in Table 5 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁻; and

(c) classifying said breast cancer patient as having a good prognosis or a poor prognosis on the basis of the level of expression of said first plurality of genes,

wherein said breast cancer patient is “ER/AGE high” if the ratio of the log₁₀(ratio) of ER gene expression to age exceeds a predetermined value, and “ER/AGE low” if the ratio of the log₁₀(ratio) of ER gene expression to age does not exceed said predetermined value.

13. The method of claim 12, wherein said control is the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁻, *BRCA1* individuals, if said breast cancer patient is ER⁻, *BRCA1*; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁻, sporadic individuals if said breast cancer patient is ER⁻, sporadic; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE high individuals, if said breast cancer patient is ER⁺, ER/AGE high; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE low, LN⁺ individuals where said breast cancer patient is ER⁺, ER/AGE low, LN⁺; or the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE low, LN⁻ individuals where said breast cancer patient is ER⁺, ER/AGE low, LN⁻.

14. The method of claim 13, wherein each of said individuals has a poor prognosis.

15. The method of claim 13, wherein each of said individuals has a good prognosis.

16. The method of claim 14, wherein said classifying step (c) is carried out by a method comprising comparing the level of expression of each of said plurality of genes or their encoded proteins in a sample from said breast cancer patient to said control, and classifying said breast cancer patient as having a poor prognosis if said level of expression

correlates with said average level of expression of the corresponding genes or their encoded proteins in said control more strongly than would be expected by chance.

17. The method of claim 12, wherein said predetermined value of ER is calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said individual.

18. The method of claim 12, wherein said individual is ER⁻, *BRCAl*, and said plurality of genes comprises two of the genes for which markers are listed in Table 1.

19. The method of claim 12, wherein said individual is ER⁻, *BRCAl*, and said plurality of genes comprises all of the genes for which markers are listed in Table 1.

20. The method of claim 12, wherein said individual is ER⁻, sporadic, and said plurality of genes comprises two of the genes for which markers are listed in Table 2.

21. The method of claim 12, wherein said individual is ER⁻, sporadic, and said plurality of genes comprises all of the genes for which markers are listed in Table 2.

22. The method of claim 12, wherein said individual is ER⁺, ER/AGE high, and said plurality of genes comprises two of the genes for which markers are listed in Table 3.

23. The method of claim 12, wherein said individual is ER⁺, ER/AGE high, and said plurality of genes comprises all of the genes for which markers are listed in Table 3.

24. The method of claim 12, wherein said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises two of the genes for which markers are listed in Table 4.

25. The method of claim 12, wherein said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

26. The method of claim 12, wherein said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises two of the genes for which markers are listed in Table 4.

27. The method of claim 12, wherein said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

28. The method of claim 12, further comprising determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

29. A method for assigning an individual to one of a plurality of categories in a clinical trial, comprising:

(a) classifying said individual as ER⁻, *BRCAl*, ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻;

(b) determining for said individual the level of expression of at least two genes for which markers are listed in Table 1 if said individual is classified as ER⁻, *BRCAl*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻;

(c) determining whether said individual has a pattern of expression of said at least two genes that correlates with a good prognosis or a poor prognosis; and

(d) assigning said individual to one category in a clinical trial if said individual has a good prognosis, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis.

30. The method of claim 29, wherein said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a).

31. The method of claim 29, wherein said individual is additionally assigned to a category in said clinical trial on the basis of any other clinical, phenotypic or genotypic characteristic of breast cancer.

32. The method of claim 29, further comprising determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis of breast cancer, and determining from the expression of said second plurality of

genes, in addition to said first plurality of genes, whether said individual has a good prognosis or a poor prognosis.

33. A method of identifying a set of genes informative for a condition, said condition having a plurality of phenotypic or genotypic characteristics such that samples may be categorized by at least one of said phenotypic or genotypic characteristics into at least one characteristic class, said method comprising:

- (a) selecting a plurality of samples from individuals having said condition;
- (b) identifying a first set of genes informative for said characteristic class using said plurality of samples;
- (c) predicting the characteristic class of each of said plurality of samples;
- (d) discarding samples for which said characteristic class is incorrectly predicted;
- (e) repeating steps (c) and (d) at least once; and
- (f) identifying a second set of genes informative for said characteristic class using samples in said plurality of samples remaining after step (e).

34. The method of claim 6, wherein said cancer is breast cancer.

35. A method for assigning an individual to one of a plurality of categories in a clinical trial, comprising:

- (a) classifying the individual into one of a plurality of condition categories differentiated by at least one genotypic or phenotypic characteristic of the condition;
- (b) determining the level of expression, in a sample derived from said individual, of a plurality of genes informative for said condition category;
- (c) determining whether said level of expression of said plurality of genes indicates that the individual has a good prognosis or a poor prognosis; and
- (d) assigning the individual to a category in a clinical trial on the basis of prognosis.

36. A method for identifying one or more sets of informative genes or markers for a condition in an organism, comprising:

(a) subdividing a plurality of individuals or samples derived therefrom of said organism subject to said condition into a plurality of classes based on one or more clinical, phenotypic or genotypic characteristics of said organism, wherein each said class consists of a plurality of individuals or samples derived therefrom of said organism each having said one or more clinical, phenotypic or genotypic characteristics specific for said class; and

(b) attempting to identify for each of one or more of said plurality of classes a set of genes or markers informative for said condition in individuals in said class,

wherein, if a set of genes or markers informative for said condition in individuals in said class is obtained for any of said one or more of said plurality of classes, said set of genes or markers is taken as a set of informative genes or markers for said condition in said organism.

37. The method of claim 36, further comprising, for each of one or more of said classes in which a set of genes or markers informative for said condition in individuals in said class cannot be obtained, repeating said steps (a) and (b) on said plurality of individuals or samples derived therefrom in said class such that said plurality of individuals or samples derived therefrom in said class is subdivided into a plurality of additional classes based on one or more clinical, phenotypic or genotypic characteristics of said organism which are different from those used for defining said class, wherein, for each of said plurality of additional classes, if a set of genes or markers informative for said condition in individuals in said class is obtained, said set of genes or markers is taken as a set of informative genes or markers for said condition in said organism.

38. A method for identifying one or more sets of informative genes or markers for a condition in an organism, comprising:

(a) subdividing a plurality of individuals or samples derived therefrom of said organism subject to said condition into a plurality of classes based on one or more clinical, phenotypic or genotypic characteristics of said organism, wherein each said class consists of a plurality of individuals or samples derived therefrom of said organism each having said one or more clinical, phenotypic or genotypic characteristics specific for said class;

(b) attempting to identify for each of one or more of said plurality of classes a set of genes or markers informative for said condition in individuals in said class, wherein if a set of genes or markers informative for said condition in individuals in said class is identified for any of said one or more of said classes, said set of genes or markers is taken as a set of informative genes or markers for a condition in said organism; and

(c) for each of one or more of said classes in which a set of genes or markers informative for said condition in individuals in said class cannot be obtained, repeating said steps (a) and (b) on said plurality of individuals or samples derived therefrom in said class such that said plurality of samples or individuals in said class is subdivided into a plurality of additional classes based on one or more clinical, phenotypic or genotypic characteristics of said organism which are different from those used those used for defining said class, wherein, for each of one or more of said plurality of additional classes, if a set of genes or markers informative for said condition in individuals in said class is obtained, said set of genes or markers is taken as a set of informative genes or markers for a condition in said organism.

39. The method of claim 38, wherein said condition is a type of cancer, and wherein each of said sets of genes or markers is informative of prognosis of individuals in a corresponding class.

40. The method of claim 39, wherein said condition is breast cancer, and wherein said one or more clinical, phenotypic or genotypic characteristics comprises age, ER level, ER/AGE, BRAC1 status, and lymph node status.

41. The method of claim 39, further comprising generating a template profile comprising measurements of levels of genes or markers of said set for said class representative of levels of the genes or markers in a plurality of patients having a chosen prognosis level.

42. A method for predicting a breast cancer patient as having a good prognosis or a poor prognosis, comprising:

(a) classifying said breast cancer patient into one of the following classes: (a1) ER⁻, *BRCA1*; (a2) ER⁻, sporadic; (a3) ER⁺, ER/AGE high; (a4) ER⁺, ER/AGE low, LN⁺; or (a5) ER⁺, ER/AGE low, LN⁻;

(b) determining a profile comprising measurements of a plurality of genes or markers in a cell sample taken from said breast cancer patient, said plurality of genes markers comprising at least two of the genes or markers corresponding to the markers in (b1) Table 1 if said breast cancer patient is classified as ER⁻, *BRCA1*; (b2) Table 2 if said breast cancer patient is classified as ER⁻ sporadic; (b3) Table 3 if said breast cancer patient is classified as ER⁺, ER/AGE high; (b4) Table 4 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁺; or (b5) Table 5 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁻; and

(c) classifying said breast cancer patient as having a good prognosis or a poor prognosis based on said profile of said plurality of genes or markers,

wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, wherein said ER/AGE is a metric of said ER level relative to the age of said patient, and wherein LN⁺ designates a greater than 0 lymph nodes status in said patient and LN⁻ designates a 0 lymph nodes status in said patient.

43. The method of claim 42, wherein step (c) is carried out by a method comprising comparing said profile to a good prognosis template and/or a poor prognosis template, and wherein said patient is classified as having a good prognosis if said profile has a high similarity to a good prognosis template or has a low similarity to a poor prognosis template or as having a poor prognosis if said profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of genes or markers representative of levels of said genes or markers in a plurality of good outcome patients and said poor prognosis template comprising measurements of said plurality of genes or markers representative of levels of said genes or markers in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-reoccurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a patient who has reoccurrence of metastases within a second period of time after initial diagnosis.

44. The method of claim 43, further comprising determining said profile, said ER level, said LN status, and/or, said ER/AGE.

45. The method of claim 44, wherein said profile is an expression profile comprising measurements of a plurality of transcripts in a sample derived from said patient, wherein said

good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

46. The method of claim 45, wherein said expression profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample.

47. The method of claim 43, wherein said profile comprises measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome patients.

48. The method of claim 46, wherein measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

49. The method of claim 48, wherein similarity of said expression profile to said good prognosis template is represented by a correlation coefficient between said expression profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

50. The method of claim 48, wherein similarity of said expression profile to said good prognosis template is represented by a distance between said cellular constituent profile and said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance equal to or greater than said given value indicates a low similarity.

51. The method of claim 49, wherein said correlation threshold is 0.5.

52. The method of claim 51, wherein said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene in said control sample, and wherein said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and wherein said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65.

53. The method of claim 52, wherein said gene encoding said estrogen receptor is the estrogen receptor α gene.

54. The method of claim 53, wherein said ER/AGE is classified as high if said ER level is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

55. The method of claim 54, wherein said estrogen receptor level is measured by a polynucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, wherein said control sample is a pool of breast cancer cells of different patients, and wherein $c = 0.1$ and $d = 42.5$.

56. The method of claim 55, wherein said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients.

57. The method of claim 55, wherein said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript of said gene encoding said estrogen receptor.

58. The method of claim 42, wherein said individual is ER⁻, *BRCA1*, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 1.

59. The method of claim 42, wherein said individual is ER⁻, *BRCA1*, and said plurality of genes comprises all of the genes for which markers are listed in Table 1.

60. The method of claim 42, wherein said individual is ER⁻, sporadic, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 2.

61. The method of claim 42, wherein said individual is ER⁻, sporadic, and said plurality of genes comprises all of the genes for which markers are listed in Table 2.

62. The method of claim 42, wherein said individual is ER+, ER/AGE high, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 3.

63. The method of claim 42, wherein said individual is ER+, ER/AGE high, and said plurality of genes comprises all of the genes for which markers are listed in Table 3.

64. The method of claim 42, wherein said individual is ER+, ER/AGE low, LN+, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 4.

65. The method of claim 42, wherein said individual is ER+, ER/AGE low, LN+, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

66. The method of claim 42, wherein said individual is ER+, ER/AGE low, LN⁻, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 4.

67. The method of claim 42, wherein said individual is ER+, ER/AGE low, LN⁻, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

68. The method of claim 42, wherein said profile further comprises one or more genes for which markers are not found in Tables 1-5, wherein said one or more genes are informative for prognosis.

69. A method for assigning an individual to one of a plurality of categories in a clinical trial, comprising assigning said individual to one category in a clinical trial if said individual has a good prognosis as determined by the method of any one of claims 7-33, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis as determined by the method of any one of claims 7-33.

70. The method of claim 69, wherein said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a).

71. The method of claim 69, wherein said individual is additionally assigned to a category in said clinical trial on the basis of one or more other clinical, phenotypic or genotypic characteristic of breast cancer.

72. The method of claim 69, further comprising determining in said cell sample the levels of expression of said one or more genes for which markers are not found in Tables 1-5, and determining from said expression levels of said one or more genes, whether said individual has a good prognosis or a poor prognosis.

73. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in any one of Tables 1-5.

74. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 1.

75. The microarray of claim 74, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to each of the genes listed in Table 1.

76. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 2.

77. The microarray of claim 76, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to a sequence in each of the genes listed in Table 2.

78. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 3.

79. The microarray of claim 78, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to a sequence in each of the genes listed in Table 3.

80. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 4.

81. The microarray of claim 80, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to a sequence in each of the genes listed in Table 4.

82. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 5.

83. The microarray of claim 82, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to a sequence in each of the genes listed in Table 5.

84. The microarray of any of claims 73-83, wherein said plurality of polynucleotide probes constitutes at least 50% of the probes on said microarray.

85. The microarray of any of claims 73-83, wherein said plurality of polynucleotide probes constitutes at least 90% of the probes on said microarray.

86. The microarray of claim 73, wherein said plurality of polynucleotide probes comprises probes complementary and hybridizable to at least 75% of the genes listed in Table 1, Table 2, Table 3, Table 4, or Table 5, wherein said plurality of polynucleotide probes, in total, constitutes at least 50% of the probes on said microarray.

87. A kit comprising the microarray of any one of claims 73-83 in a sealed container.

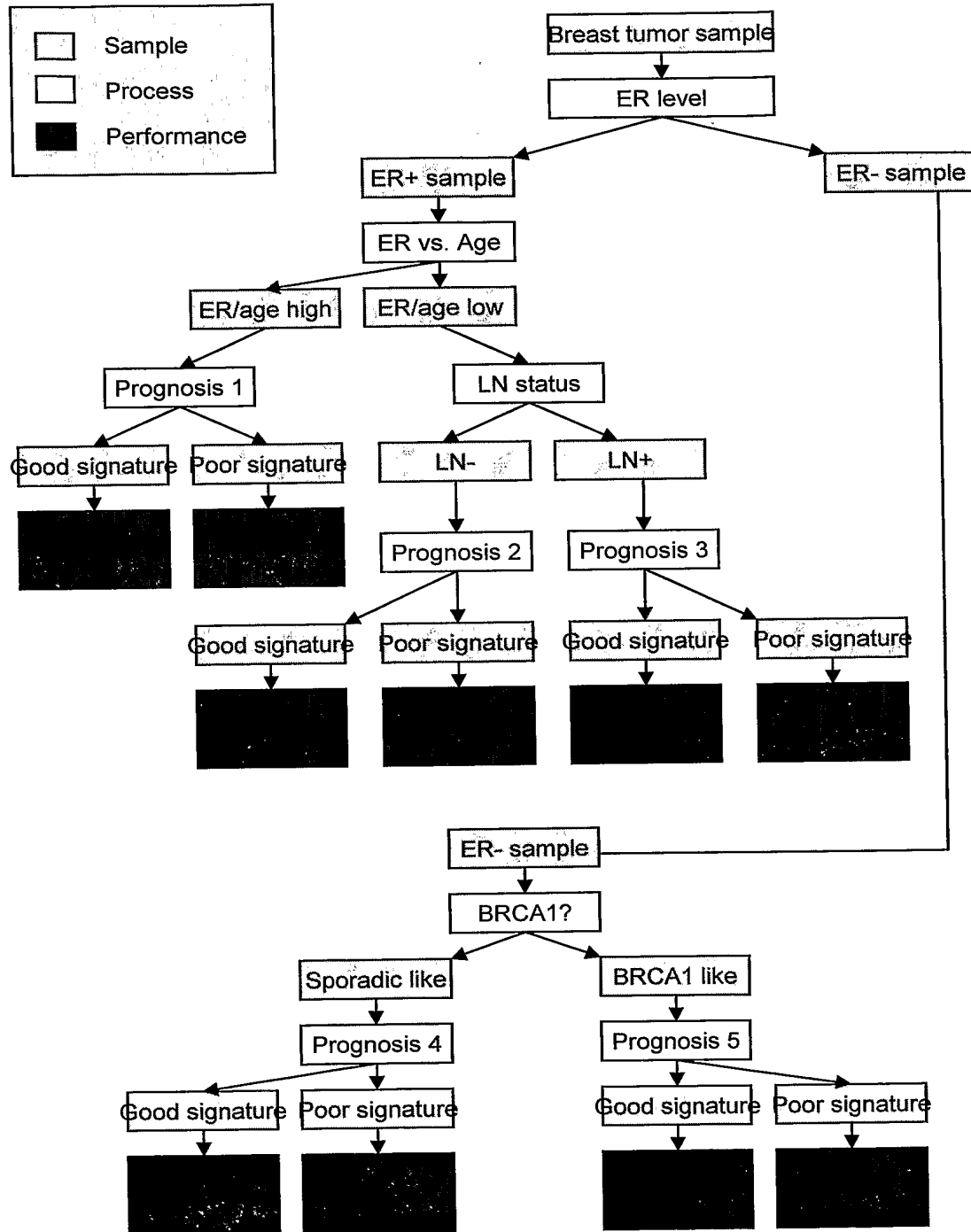


FIG. 1

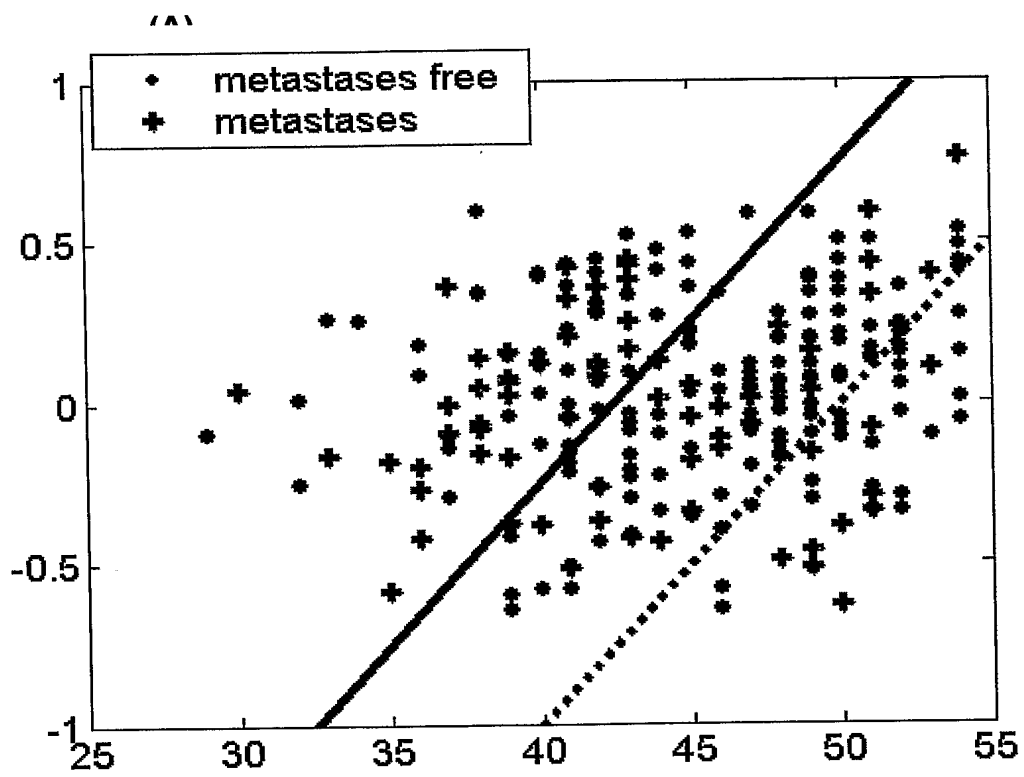
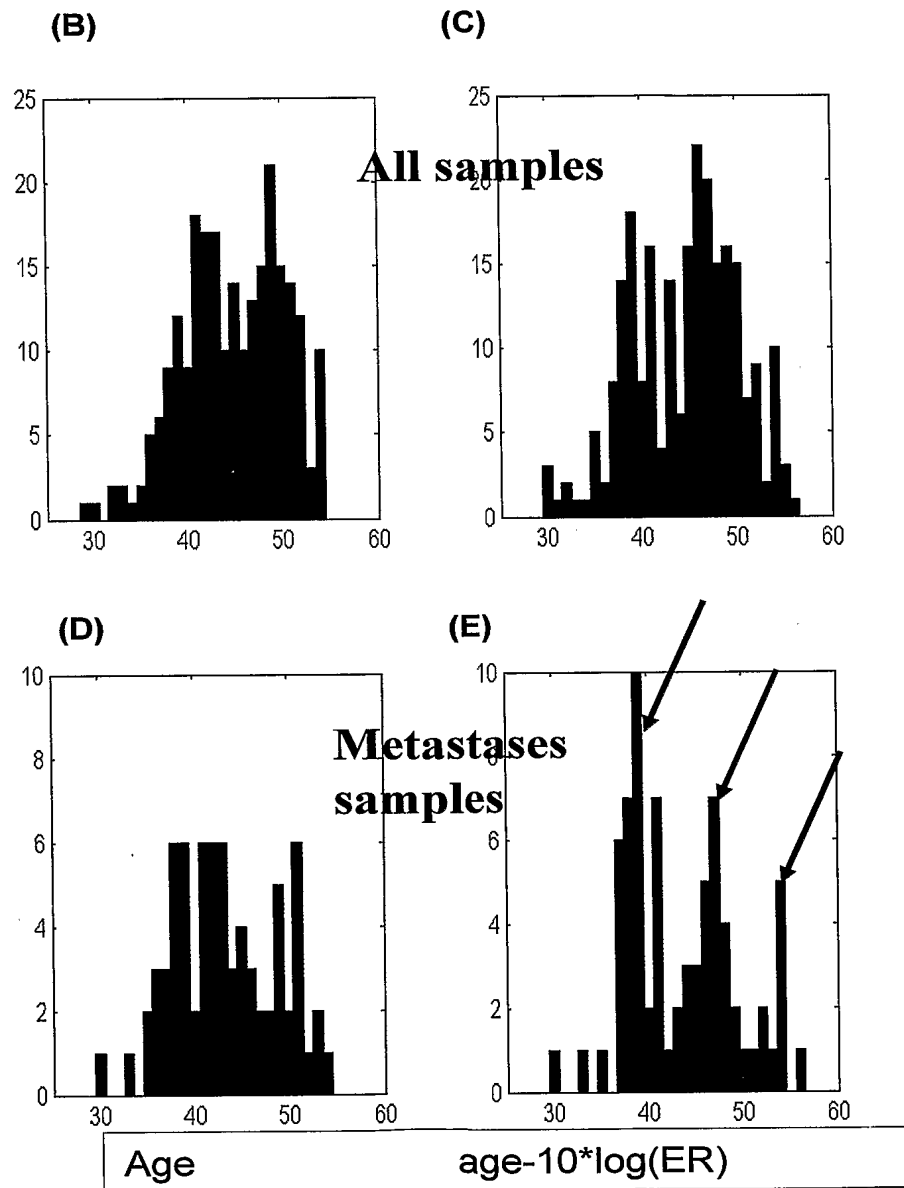
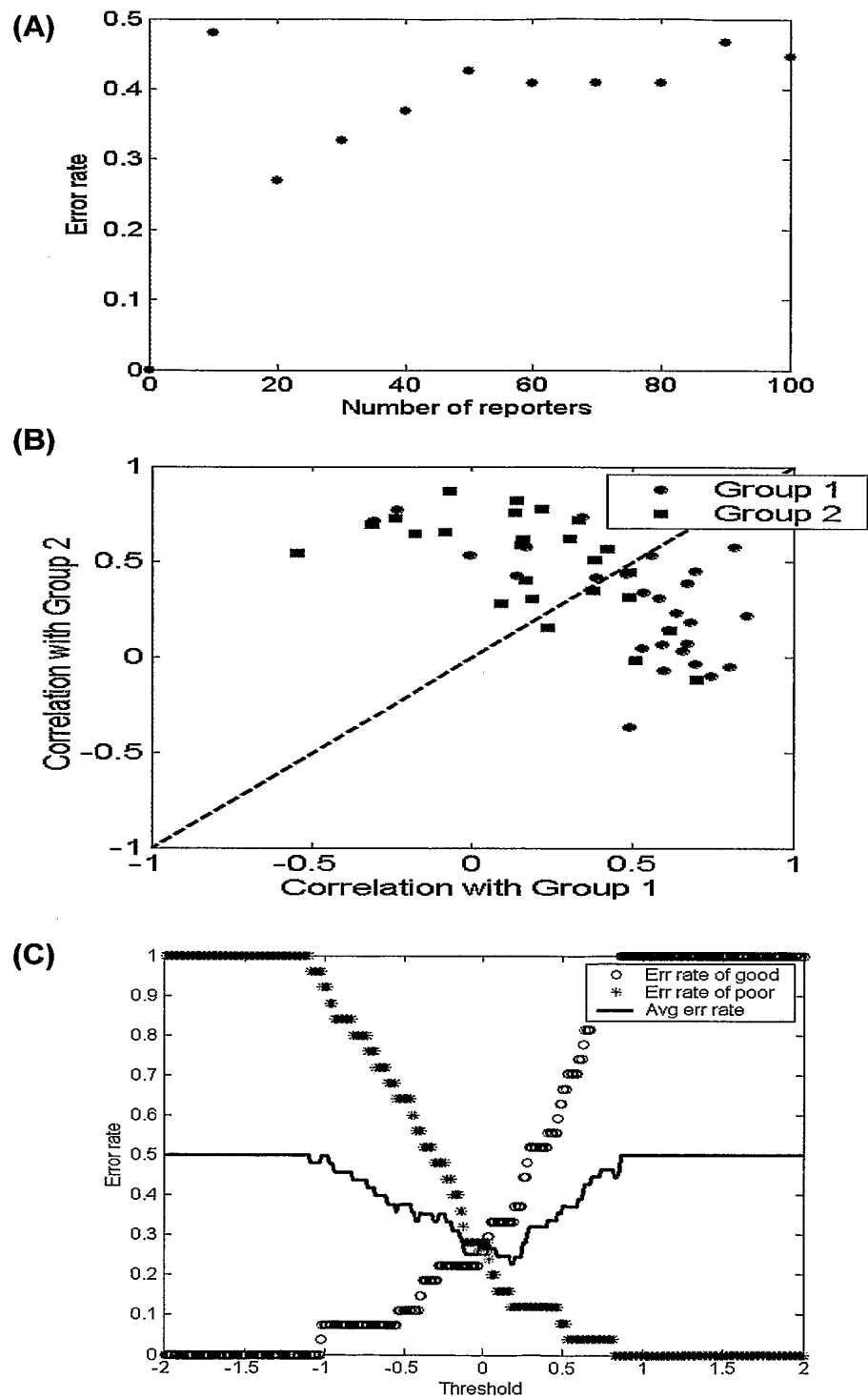


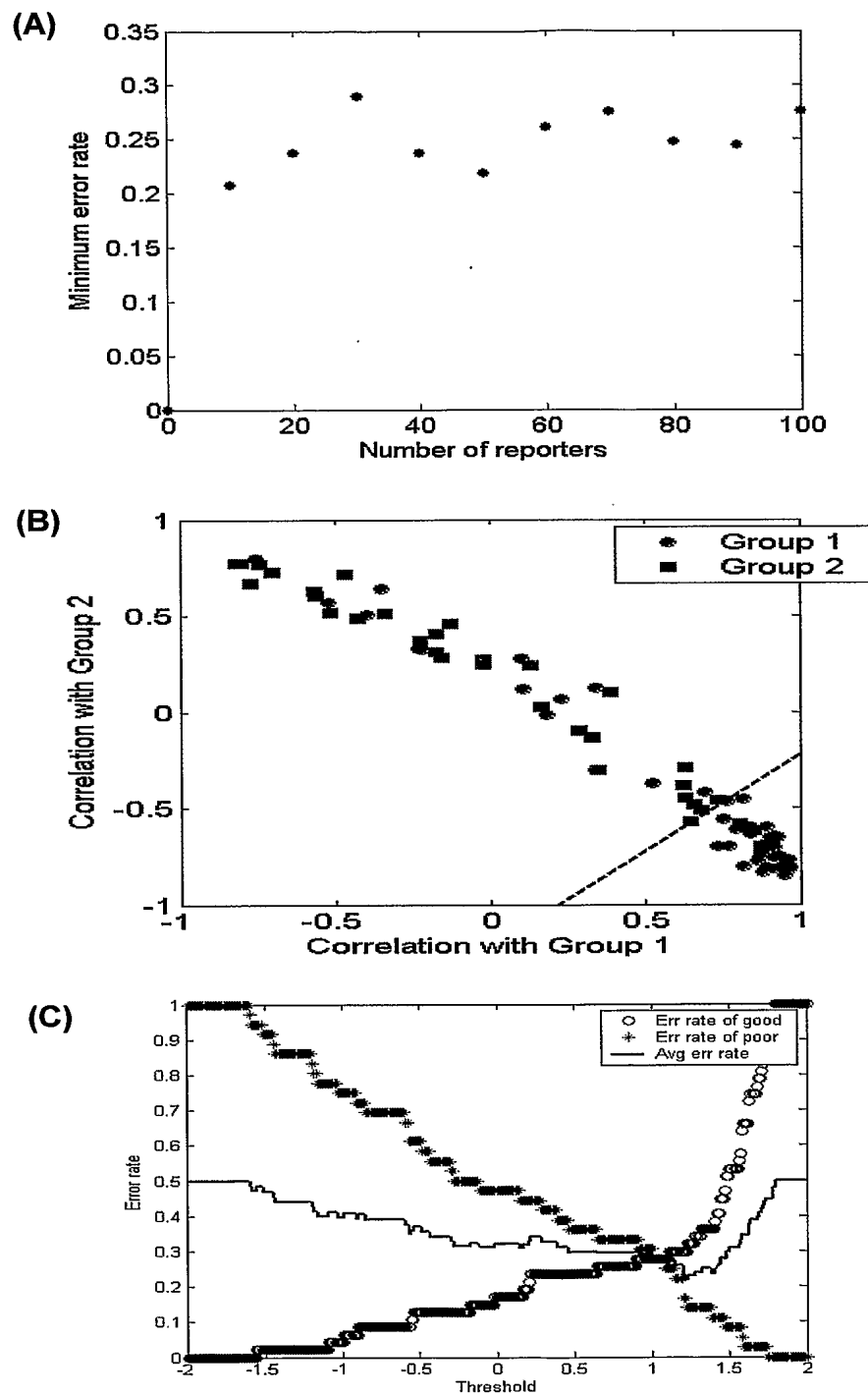
FIG. 2A



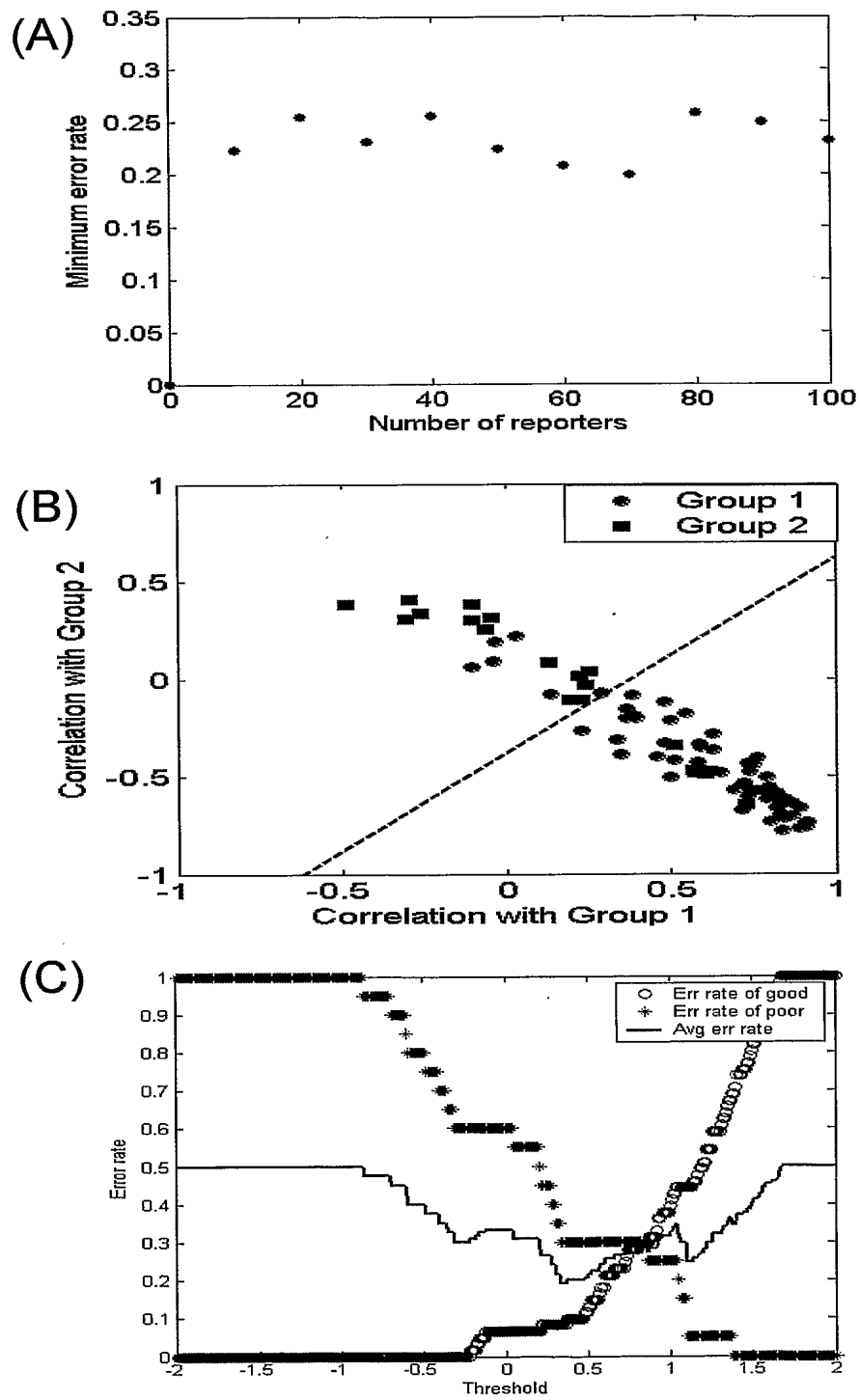
FIGS. 2B-E



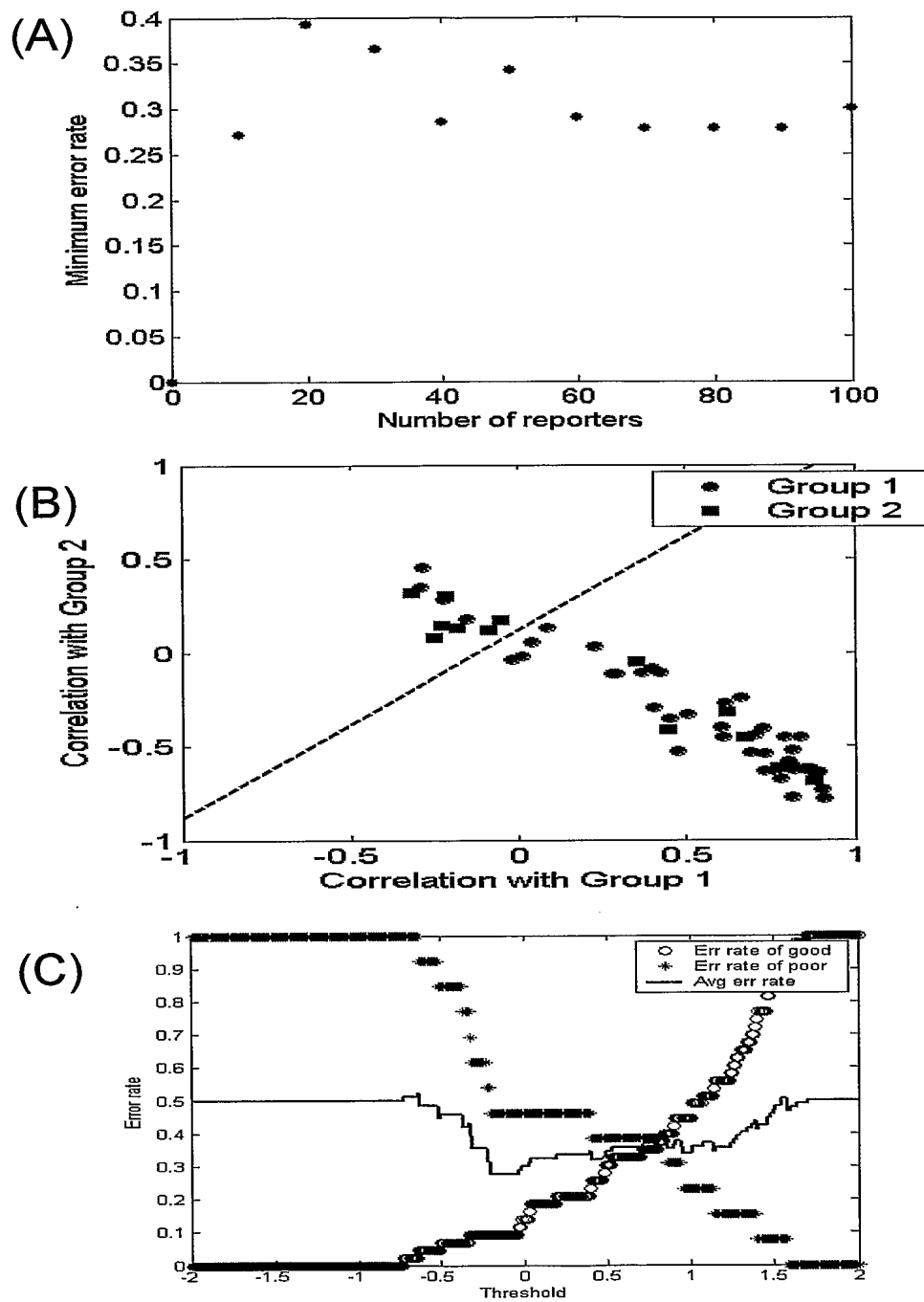
FIGS. 3A-C



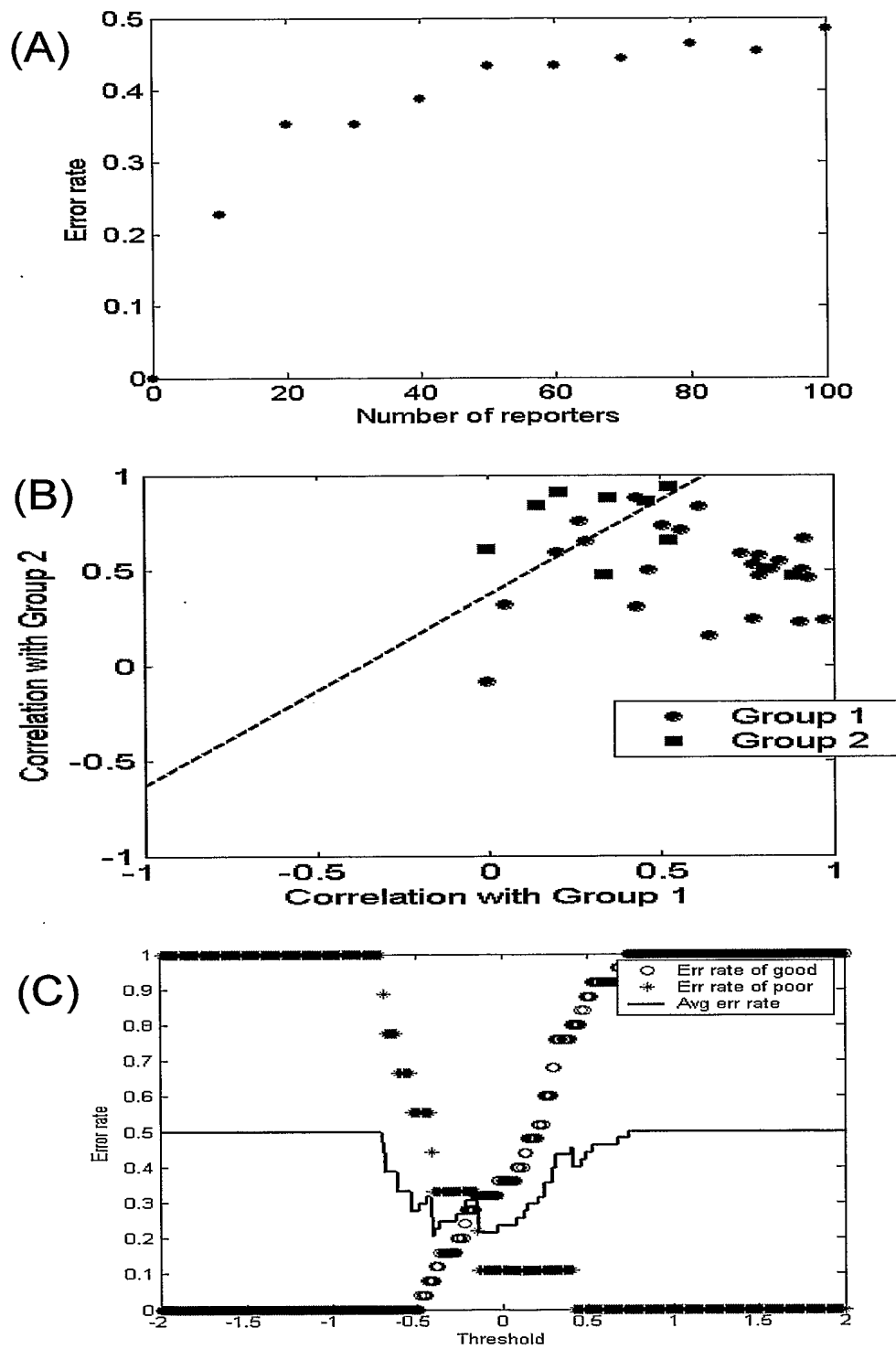
FIGS. 4A-C



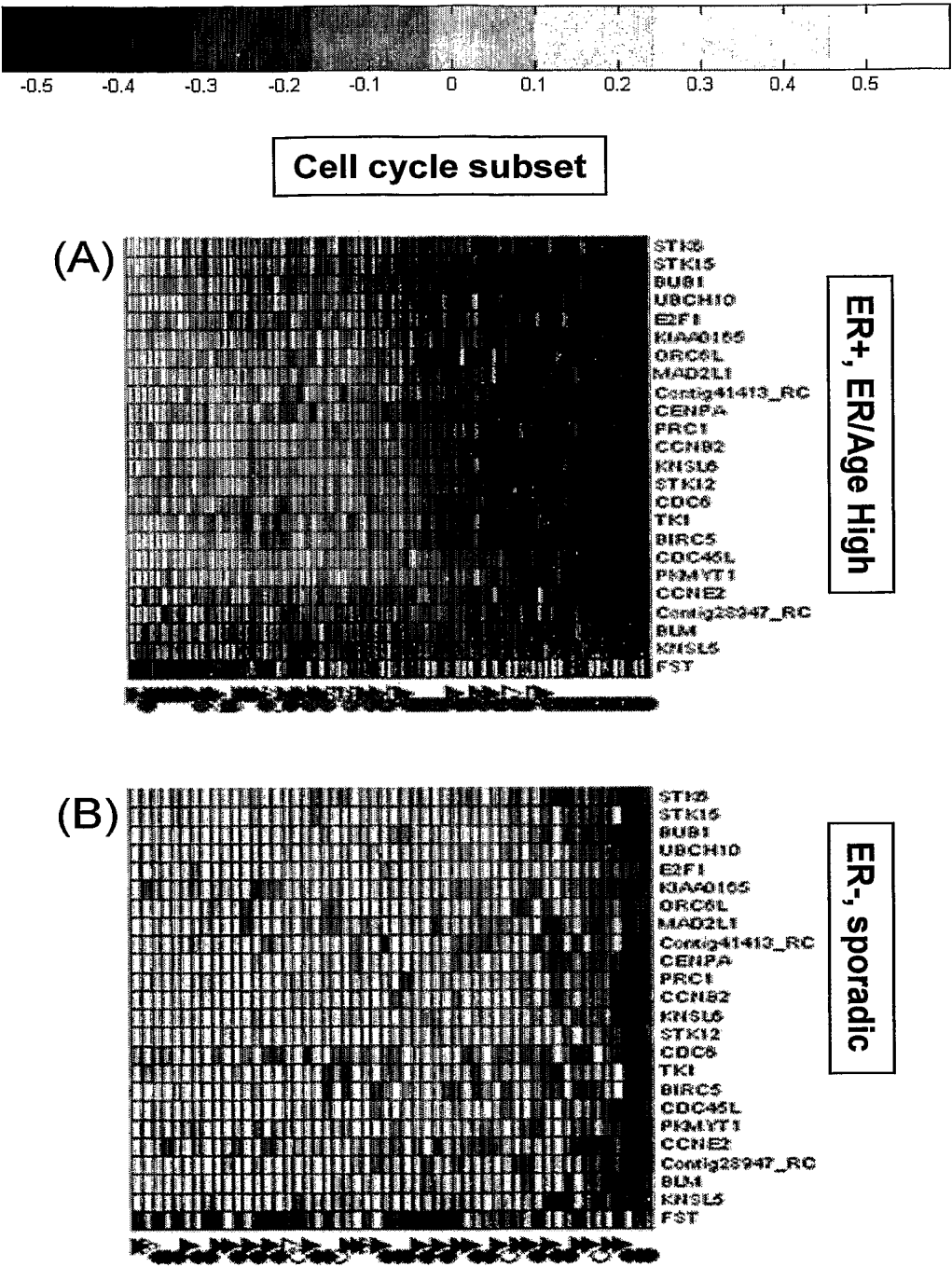
FIGS. 5A-C



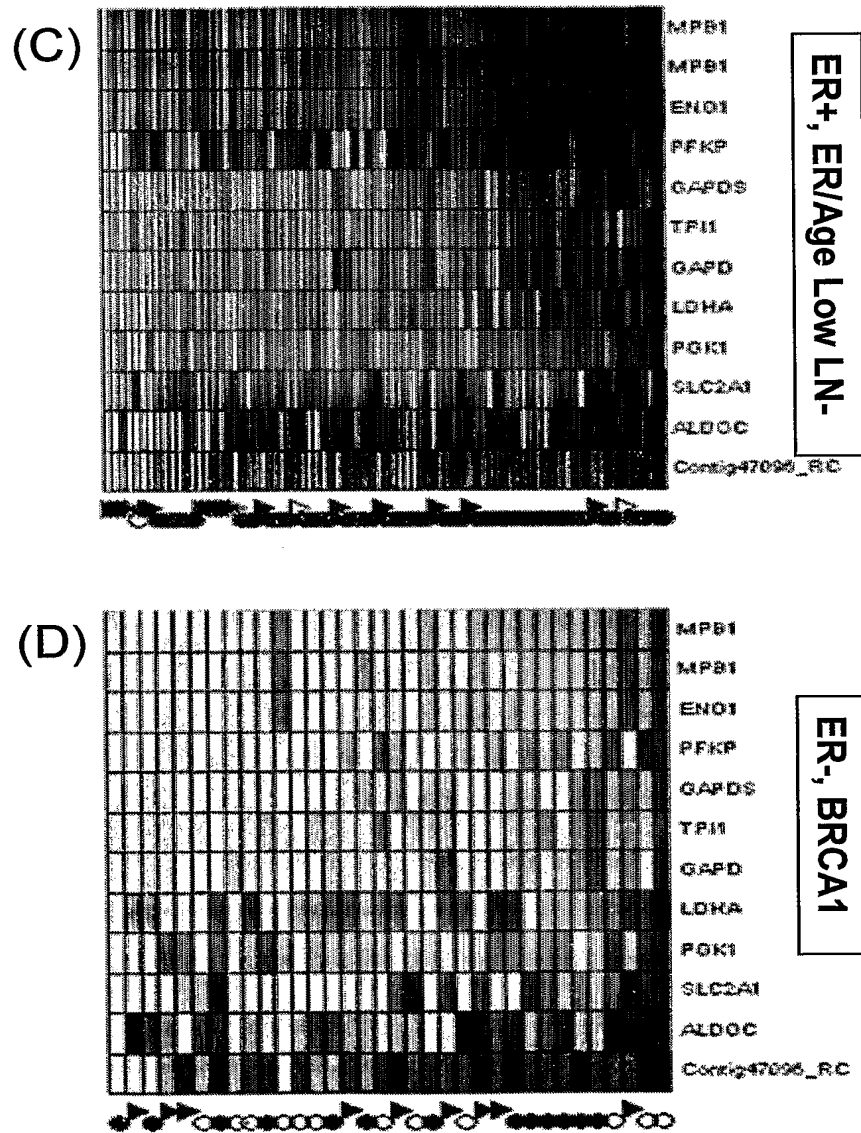
FIGS. 6A-C



FIGS. 7A-C



FIGS. 8A and 8B

Glycolysis subset

FIGS. 8C and 8D

Sequence Listing

<110> Rosetta Inpharmatics, LLC

<120> Classification of Breast Cancer Patients Using a Combination
of Clinical Criteria and Informative Genesets

<130> 9301-251-228

<140>

<141>

<150> 60/650,401

<151> 2005-02-04

<150> 60/604,076

<151> 2004-08-24

<150> 60/550,810

<151> 2004-03-05

<160> 366

<210> 1

<211> 4946

<212> DNA

<213> Homo sapiens

<300>

<308> AB032969

<400> 1

cagcctcagc	ccccagatga	agatgggggat	cacagtgaca	aagaagatga	acagcctcaa	60
gtggtggttt	taaaaaaggg	agacctgtca	gttgaagaag	tcatgaaaat	taaagcagaa	120
ataaaggctg	ccaaagcaga	tgaagaacca	actccagccg	atggaagaat	catatatcga	180
aaaccagtc	agcatccctc	agatgaaaaa	tattcagggt	taacagcaag	ctcaaaaaag	240
aagaagccaa	atgaagatga	agtaaatcag	gactcggtca	aaaagaactc	acaaaaacaa	300
attaaaaata	gtagcctcct	ttcttttgac	aacgaagatg	aaaatgagta	agtgtaaata	360
ttttgaattt	agtctacttt	gaaagtatat	ggagtgttca	ttaaaatcac	attttttcct	420
attataaaga	tactacaagt	tctttataga	aagtttagga	aatagagaaa	aaaatttaat	480
aaactacatc	tattcatcaa	taccctcttg	acttaaaatg	ccaactctat	agaaattagc	540
tagtattaac	attttgttat	ttcccttggt	tggttgtata	tatatgtaaa	ttataatttt	600
aagcaaaata	cattttttgt	gtgtaaaaca	aattttataa	atacaactgt	attgcaaattg	660
ttctttgtcc	tgcttctcac	ttgacattgc	attatgagta	ttcttccagg	tcagtaaatt	720
tcaaaaacct	gacattaata	gctacagata	atttcataaa	catctcattg	tatctttttc	780
attagcaata	gctccacttt	gggtggggga	gatgataatg	tgcccttgta	aaaataacctc	840
cccaactcct	gctaaggggtg	gccatgagac	tcagctctgg	caagttaaga	aatacagggtg	900
gaattctgct	tgataaagct	gctgggtttt	ttgttacaaa	aggacagact	tggcaaacat	960
gagcctttgc	tcttatcttt	tcctctact	tggtgtgcag	agataaaacc	tgagtaccag	1020
agccactttt	aggcataagg	aaggcagcca	tggtgctttg	gtcatgttag	taaaaagact	1080
cagagcttgg	ctccttgctg	acatgcctgg	aggagctgct	acaccagctt	ggattgctga	1140
cctctgactt	cttggtagtg	agaagaataa	acactgtgct	taattaggcc	ttggtcaggt	1200
ttcttttata	tgacagccaa	tgagtcctca	agtaatacaa	taaataactg	gtcaaaactgt	1260
tactggtgga	gggtgtccag	gttcttgcca	ttttggacaa	ataattgaac	aaaacgcaca	1320
aagcaatgaa	tatcctctag	aggtttgcca	ttggttactt	ggcgtacacc	ctgtgtaaat	1380
gaagtagtgg	cccgtagcct	gtctgattgg	tgacagaaag	gaccaatcag	aggctgaagt	1440
gaagttacaa	agttatactc	ctgtgtaaat	gaggacttgg	cctatgacca	gtctgattgg	1500
ttgcaggagg	ggaccaatca	gaggcacttt	catttttcat	ctgcaatgca	gaaaaggcaa	1560
ggggattgca	aaggagtag	cctctgatcc	ttttgttact	taggtatgga	gaggtgggggt	1620
tttccttttg	attcagttct	aggaagtcaa	tgtgaatcag	ccttaggttc	cctgtctcca	1680

```

gaccctattc tcttgcctca ttttccccct gagagacgtg atcctctgtaa atctttatgg 1740
gaggctgaga gactgagggt ctttcttctg taactgcttc atgctaactt gggacacagt 1800
ccctacctat tggagatcac gtaactctca ccttgccttg tctaggggag acagggtagc 1860
ttcttgatgg ccggtgggtg cttctcctga aactggctag aaatcttgtc acatgatcat 1920
ctaacttggg ggtctctagg caaaaggaaa tggatttggg taaaagattt aacagatatg 1980
gtccaaaaac caaggcaaat ataatcatta ataatgggct ggccaaggga gggagccatg 2040
aaaccaactt tagtgccctt taggtgcccc agctgttgtc atattttaga ggcccagtca 2100
gctagttttc aggtgggtgc ccttactaat cctgattggg tgacatcaaa acagcattct 2160
tcttctagga aaatacataa gccacctggt tcagcagtta ggagatctag tccccctcga 2220
ttttgcaaag cgaccactgc caaggagcct atccgaattt gtaagggtgac aatactttga 2280
gcaatgttat ccaggccttc cataaaatcc ttggacaagc gttggtaata ggatagggaa 2340
gttgcaatcc cgctaactcc cattcctacc tctgctgtta ttctagccg ttgtgtctgg 2400
tggttgcatg taaaggata atgagggatt ggttgttggg agctatatta atttagggac 2460
atacaatatt tctgtctcca gtctaccact tccaccaag acaaatcaca gcagaaccga 2520
cctaacttca aaataaactg cagtcccata tactgggcct gattaccacac acaaagtga 2580
acaagaatca ttgtccatat agactctcct agattggcct tgctagaaca tttcacaagg 2640
ccatttcagt caaagtcctg agaaagtaac cggtttcaat tgtgccctat tacaaaagaa 2700
aacgtgggta ttaactttat acagacaaat gccatgaatt aagaatatc ataaatagtt 2760
tacaaattct ggagaaatta gaatactcaa tacacttaaa gtgtatttca aggctataaa 2820
tagctcaaaa taaaaagatt attcagactc tgaaaaaaca aaaagaagta gcaatatttc 2880
aaacaacaaa attatctcag tcttccatta ttctccatta gttcatttca gtccatgtaa 2940
tcaactcctg ctctacttca tattcatctt tatgaacaca tcagccttct aattagtgcc 3000
ttggaagttt tctgtctaat ccaatggcac actctccaaa gttaccagaa acctgcattc 3060
aagagttctt ttcatgaact ccaaagaagt aagccttggg ctgtagctga ttataagtc 3120
cttttttttt ttgagaagga tcaaagcaaa acatcaatta tggatgacaa aagtcttaag 3180
acagccataa agacacagtt gacaaatgtg gctatttctg tggcttacia caatttaaca 3240
taatcattac aacatatatt agaatttttag aactctcata caatcctgga 3300
acacatatta acaacaaatc tctatcagta taaccctaaag gaagctaaac accacctcac 3360
acttgacaat gtttctctga taattcaaac attacaaata agcctaatat aagcctaata 3420
tgtcactctt gaacttcagg aagcctaata tccaaaaagt tagtttaagg tcaaaagttt 3480
ttgaattaac ttttttccat tagtatgggc atatctttct tactaatttg taagttatgt 3540
aatttatcaa tttttttttg ttgttctgtt tcccaacctc tatgtcagat aaagaatcac 3600
ccaggccaga ccagctggct catgcttggt gtcccagcac tttgggaagc caaggtggga 3660
gaattgcctg aagccaggaa tctgagcca gcctgggcga caaagcaata cccctatctc 3720
tacaaaaaat aaaaaatagc caggtgtggc gacacacacc tgtggtccca gctgctcggg 3780
aggctgagcg ggaggatggc ttgggccagc gggttcaacg ctgcagtga ctgtgattgc 3840
gccactgcac tccagcctgg gcaacagagt aagaactgtc tcaaaaaaaa taaaaaatag 3900
aaataaattt taaaaaaga attaccata ttctctttgt ttttgtttat tcacattaac 3960
ctttattcta tctggaattt atttgagta acttttttct caaataatca attgtcctag 4020
aaccatgtgt ttctcattta tttgaaagc catctagtga gagatttctc caaatgttgg 4080
ggtaggggaag ggaggggaag cactttaaag tctgagcctt tagaggtgat tctcaagac 4140
cctgcttaat cctaacaatt ttctcatta gtaaaagtca gcccaaactg ggggcttggt 4200
aagatcctta ccagccacat ccactgaaa ttatgaattt caaagtatct tacaaatttg 4260
gtgccacatt atctttttta agtttgtttt gttttgtttt tttgagacag agtctcgctc 4320
tgtcaccggg gctggagtgc agtggcgcca tctcagctca ctgcaagctc cgcctcctgg 4380
gttcacacca ttctcttggc tcggcctccc aagatctctg gactgcagtc gcccgccacc 4440
acgcccggct aatttttttg tatttttaat agagacgggg tttcactttg ttagccagga 4500
tggtctcaat ctctgacct catgatccac ctgcctcggc ctcccaaagt gctgggatta 4560
caggcaggag ccaccgcgcc tgggcctttt tttaaagttt aagtacctat aaagaacact 4620
gaaagtgat gtgtgtggat gagctaggaa gacctgaaat aggtctctct taaattaatc 4680
aaattaatcc tgaagccatt ctgcaatact gtctttaatg tatactcact tgttatagaa 4740
gccagggttt tttcccctaa tttgtatcat tgctatatgt gttattgtac caaatacac 4800
tgtttttaatt gctgtaaaatt ttaatatgtc ttagtatctg ggtgtgggaa tcttgaaagc 4860
atggagtttg tgttattcac cactgtattc tcaaatatca gaagagtatc tggcctacta 4920
agtgcacaaat aaacatagtt aaaatg 4946

```

<210> 2

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AB032969

<400> 2

taatcctgaa gccattctgc aatactgtct ttaatgtata ctcaattggt atagaagcca 60

<210> 3

<211> 1007

<212> DNA

<213> Homo sapiens

<300>

<308> AF005487

<400> 3

gaatacagaa	tgtggggcaaa	ctcgtctctg	tgccggccgc	cagaagggtt	gctgagggca	60
atcactccct	ggtgccgggc	tccttgaggt	tatgcactgg	gacatctaga	gcctattggt	120
tgaggaatgc	agtcttgcaa	gcctgctctg	gatcaagcca	cagactgaaa	cacccccgaa	180
gagcaagcac	gtttcttgga	gcaggctaag	tgtgagtgtc	atatcttcaa	tgggatgaag	240
cgggtgcagt	acctgaacag	atacatccat	aaacgggagg	agaacctgcg	cttcgacagc	300
aacgtggagg	agttccaggc	agttacggaa	ctggggcggc	ctgtcgcaga	gaactggaac	360
agccagaagg	gcatcccga	ggagaagcgg	gacaagatgg	acgactactg	cagatacaat	420
tacgggggtt	tttgagagct	tcacagtgc	gccgcgagtc	catcctaagg	tgactgtgta	480
tcctgcaaag	accagcccc	tgcatcaccg	caacccctg	gtcggctctg	tgagtgggtt	540
ctatccaggc	agcattaaag	tcagggtggt	ccagaatgg	caggaagaga	aggctgcggt	600
ggtctccata	ggcctgatcc	agaatggaga	ttggaccttc	cagacctgg	tgatgctgga	660
aacagttcct	cggagtggag	aggtttacac	ctgccaaagt	gagcatccaa	gcgtgacgag	720
ccctctcaca	gtggaatgga	gtacacggac	tgaatctgca	cagagcaaga	tgctgagtgg	780
agtcgggggc	tttgtgctgg	gcctgctctt	ccttgggaca	gggctgttca	tctacttcag	840
gaatcagaaa	ggacactctg	gacttcagcc	aacaggactc	ctgcgctgga	ctcctgagct	900
gaagtgcaca	tgaccacatt	caaggaagaa	ccttctgcca	cagctttgca	ggatgaaaag	960
ctttcccact	tggtctttat	tcttccacaa	gagctctctc	aggacca		1007

<210> 4

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF005487

<400> 4

tttgcaggat gaaaagcttt cccacttggc tcttattctt ccacaagagc tctctcagga 60

<210> 5

<211> 3200

<212> DNA

<213> Homo sapiens

<300>

<308> AF026941

<400> 5

caggaagggc	catgaagatt	aataaagatt	tggactcagg	gcaaataatt	acttagtagc	60
aataactcaa	agaattactg	ttgaataaat	aagccaatta	agcagccaat	cacgtactat	120
gcggatgcac	acaaatgaaa	ccctcacttc	aacctgaaga	cattcgcaca	tgagttacgt	180
agagggacct	gcaggaagcg	gtagagaaaa	cataaggctt	atgcgtttta	tttccacacc	240
aatttcagga	tctttgtcac	tgacagcagc	actaagactt	gttaacttta	tatagttaag	300
aagaacaagg	ctgagcgcca	tgactcacgc	ctgtaagcct	agaactttgg	gaggccaaag	360
caggcagact	gcttgagccc	aggagttcca	gaccagcctg	ggcaacatgg	caacacccca	420

```

tctctacaaa aaaatacaag aatcagctgg gcgtgggtgat gtgttccctgt aatctcagct 480
actcggggagg cagaggcagg aggtattgctt gaaccccgga ggcagagggt gtagttagcc 540
gagatctcgc cactgcactc cagtctggac gacagagtga gactcagctt caaataaata 600
aataaataca taaatataag gaaaaaata aagctgcttt ctctcttcc tctcttttgg 660
tctcatctgg ctctgctcca ggcatctgcc acaatgtggg tgettacacc tctctttt 720
gctgggaagt tcttgagtgt gttcaggcaa cctctgagct ctctgtggag gagcctggtc 780
ccgctgttct gctggctgag ggcaaccttc tggctgctag ctaccaagag gagaaagcag 840
cagctgggtcc tgagagggcc agatgagacc aaagaggagg aagaggaccc tctctgccc 900
accaccccaa ccagcgtcaa ctatcacttc actcgccagt gcaactacaa atgcggcttc 960
tgtttccaca cagccaaaac atcctttgtg ctgccccttg aggaagcaaa gagaggattg 1020
cttttgctta aggaagctgg tatggagaag atcaactttt cagggtggaga gccatttctt 1080
caagaccggg gagaatacct gggcaagttg gtgaggttct gcaaagtaga gttgcggctg 1140
cccagcgtga gcatcgtgag caatggaagc ctgatccggg agaggtggtt ccagaattat 1200
ggtgagtatt tggacattct cgctatctcc tgtgacagct ttgacgagga agtcaatgtc 1260
cttattggcc gtggccaagg aaagaagaac catgtggaaa acctcaaaa gctgaggagg 1320
tggtgtaggg attatagaat ccttttcaag ataaattctg tcattaatcg tttcaacgtg 1380
gaagaggaca tgacggaaca gatcaaaagc ctaaacctg tccgctggaa agtgttccag 1440
tgctcttaa ttgaaggtga gaattgtgga gaagatgctc taagagaagc agaaagattt 1500
gttattgggt atgaagaatt tgaaagattc ttggagcgcc acaaagaagt gtcctgcttg 1560
gtgctgaat ctaaccagaa gatgaaagac tcctacctta ttctggatga atatatgctc 1620
ttctgaact gtagaaaggg acggaaggac ccttccaagt ccatcctgga tgttggtgta 1680
gaagaagcta taaaattcag tggatttgat gaaaagatgt ttctgaagcg aggaggaata 1740
tacatatgga gtaaggctga tctgaagctg gattggtaga gcggaagtg gaacgagact 1800
tcaacacacc agtgggaaaa ctctagagt aactgccatt gtctgcaata ctatcccgct 1860
ggtatttccc agtggctgaa aacctgattt tctgctgcac gtggcatctg attacctgtg 1920
gtcactgaac acacgaataa cttggatagc aaatcctgag acaatggaaa accattaact 1980
ttacttcatt ggcttataac cttgttggtt ttgaaacagc acttctgttt ttgagtttgt 2040
tttagctaaa aagaaggaat acacacagga ataatgacct caaaaatgct tagataaggc 2100
ccctatacac aggacctgac atttagctca atgatgcgtt tgtaagaaat aagctctagt 2160
gatatctgtg ggggcaatat ttaatttggg tttgattttt taaaacaatg tttactgcta 2220
ttctatattt tccattttga aactatttct tgttccaggt ttgttcattt gacagagtca 2280
gtattttttg ccaaataatc agataaccag ttttcacatc tgagacatta caaagtatct 2340
gcctcaatta tttctgctgg ttataatgct tttttttttt tttgctttta tgccattgca 2400
gtcttgtaact ttttactgtg atgtacagaa atagtcaaca gatgtttcca agaacatatg 2460
atatgataat cctaccaatt ttcaagaagt ctctagaaag agataacaca tggaaagacg 2520
gcgtggtgca gccagccca cgggtgcctgt tccatgaatg ctggctacct atgtgtgtgg 2580
tacctgttgt gtccctttct cttcaaagat ccctgagcaa aacaaagata cgctttccat 2640
ttgatgatgg agttgacatg gaggcagtgc ttgcattgct ttgttcgct atcatctggc 2700
cacatgaggc tgtcaagcaa aagaatagga gtgtagttga gtagctgggt ggccctacat 2760
ttctgagaag tgacgttaca ctgggttggc ataagatatc ctaaaatcac gctggaacct 2820
tgggcaagga agaattgtgag caagagtaga gagagtgcct ggatttcatg tcagtgaagc 2880
catgtcacca tatcatattt ttgaatgaac tctgagtcag ttgaaatagg gtaccatcta 2940
ggtcagttta agaagagtca gctcagagaa agcaagcata agggaaaatg tcacgtaaac 3000
tagatcaggg aacaaaatcc tctccttggt gaaatatccc atgcagtttg ttgatacaac 3060
ttagtatctt attgcctaaa aaaaaatttc ttatcattgt ttcaaaaaag caaatcatg 3120
gaaaattttt gttgtccagg caaataaaaag gtcattttta tttaaaaaaa aaaaaaaaaa 3180
aaaaaaaaaa aaaaaggcca 3200

```

<210> 6
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF026941

<400> 6
 atttttgaat gaactctgag tcagttgaaa tagggtacca tctaggtcag ttttaagaaga 60

<210> 7
 <211> 1799

<212> DNA
<213> Homo sapiens

<300>
<308> AF035284

<400> 7
gcttgaaccg gggaggtgga gggtgcagtg agctgagatc acgccattgt actccagcct 60
gggcgacaga gcaagactcc atttcaaaaa aaaaaaaaaa aaaaaaaatc cactcatata 120
aaaggtgagc tcagctcact ggtccatttc tcagtggctt ctocatcctc atttgcaaac 180
ctcagaggga taaggcagtt gaacctgatg agcaagaatt ataacagcaa ggaaacatta 240
atgcttagaa ttctgagatc cagcacaact cagtctgtgg gagctcagct cgctgccag 300
ggataggtat gacctatgtc tgccttaggc tgcctgggaga tgccattctc cagtttcaga 360
agcaggcagg gcaaagggtca agactgtggg attgggggtct tttgggtctg aaggatcctg 420
gaaccactga ttttggttta ttccctccag ggtctaaaga gaacaagagg tgctagctct 480
taccaaaaca gatggtagag agagttgctg gctattttaa aagctctttc atctttta 540
tcacctcttc ttttcacctc ttttaaccact cctcagggaac agaactcttc taggactggg 600
ggctcttttag ctccataagc aagtgcagcag atggggacaag ttagtctttt ctccctagaa 660
acaaagggga tgcccagtggt tttccctttg cttcccaacc taaaatttca agtttaataa 720
aatagcaatt agcagaagtg accaaattgg gagataatta tcagtcatga ggaaagacac 780
agatttcggg cataaagaat gtaagggcta taagtagaaa ctttctataa cctaaatgat 840
gttatagaat tatttttgag caggagcaga aagattaaat atgatactt cactacttcta 900
aatcagaaat aggaagatta aaaccacaga acagtttgtg atttctattg ctggtagcta 960
ggatctttac tctgtccact cttgttcaag tatctaactc ttctggaaac caaataggct 1020
ttagaagaga ttatctata ttctatcag tataatacta aaatgtaact ttttaatcat 1080
ctgggtttta aaagataaac agtttagccc atctctccag agagcaaaca taggaatatg 1140
actcaggagc ctccatagggc ttatcatcag ccctcacacc cgcttcccc tccaaccac 1200
agcctttgct tccaggtggc aggattacta ctttgccctc tcagcagcat ctactctagg 1260
catattgatc atttttagaca ctggggagaag agaacctcaa actaggagga aaagacagag 1320
cctccactta gttttgggag gggatggcag acagtcaagg agatgagcgt cctaaggcat 1380
gttgggatag ggtcagatgc accaccatg gagaggtttg tcaacacaaa gacatggaag 1440
gttagagggt tgtcaacaaa aagacatgga aggttaggtt tgtcaacaca aagacatgga 1500
agattagagg tttgtcaaca caaagataca ggaagaatgg gctgcagaag atttagatgt 1560
tttccatttg ggcacatttt acttagctgg agaactaggt ttaaaacagc ctgggtagga 1620
aaattagaag caagctggat gcagtggctc atgacctgta tccaacact tttgggaggt 1680
ccaggcagga ggtacacttg ggcccaggag gtcaagcctg cagcgagctg agatcacacc 1740
actgcactcc agcctggggg gatagaacaa gacctgtct caaaaaaaaa aaaaaaaaaa 1799

<210> 8
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> AF035284

<400> 8
caaaaagaca tggaagggtta gggttgtcaa cacaaagaca tggaagatta gaggtttgtc 60

<210> 9
<211> 1380
<212> DNA
<213> Homo sapiens

<300>
<308> AF052162

<400> 9
gtcaaaggat atttatttat aggccttttt ttttttaata tagaatctga ggctgtttgg 60
gctttgactt aaatttccat caggcctctc tccagcaggt aatccctctc cttccgctgg 120
gtcccctggg gaggtgtgaa ctcaagggcc tagcccaaaa acactttttc tgcttttctt 180

```

aatccttttc cagtccectc tttttttata aacgttggca gtttgatgtt tctgtttcgg 240
cataacgtaa tccattttcac tgtagocctaa actccagttcc gaggttggat attgtttcaa 300
tgagcagggc ccgagctgga agcgcaaggc agccgcgcgc gtgcccgtcc tcccttgccc 360
tcaggccagg tccctgctgg aagcggctgc atcttctgt cagccctggg ttccatgggtg 420
actggcgctca cgcagccacc cgagtatggc tgaccttctc gcagagagag gagccgcagt 480
cttttgcttg tggaaggaga cgctgggctg tgcgggtgcgg aggggtgatga ggatgtctgg 540
tgacagccgt gcggacacca ctctctctct cagcaactgc tcccagcgcc agggtcgcgg 600
gcacatccca ctgagagcgg gggctcctgcc ccatcttaga gtcaaaggca gaggggcttc 660
caggccctgg atgggggtatt ttggtgtcac ctgaagtccc tctgacatca ccttgtttca 720
tcatttttta tgacagaatt agaaacccat ccttcaagca caataatcat cacagacttg 780
agtttgcttc ctaaagcaaa ggctccgggt ttgtttggaa aatttttttg atttctgaaa 840
tgaattgatt tttatatttg ggcatctct atagaaagt accaccaagg ccagtaagta 900
cgggaaaaaa tgtttactaa cttcctcaga gattcgtgat acgcgtttct cactgacag 960
acatttaaaa acaaccttca gctccgtttc aatcaatcac ctgcacttgt tttttagcat 1020
ggacactgcc agcaggacag acagggatgg agtaaaaccga agtcaatttc agggctcttg 1080
gcgtgttggc cacagaagaa atcctagtgc agccttttgt agctaacagt cactgatttt 1140
ataattggag aatgcgtaaa gattcatttt tcaaggagaa gagcctgcaa atggccaatg 1200
aaggaggtaa ataaactaag atattccgag ggaagggacc caggccacct ccttccgca 1260
ggtctgcaga tgaagggttt tttgaatgaa atgccactgt gcattttcag aaaaaaaaaa 1320
ctctgataaa cagactttga atggaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1380

```

<210> 10

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF052162

<400> 10

```

cagtaagtac gggaaaaaat gtttactaac ttcctcagag attcgtgata cgcgtttctc 60

```

<210> 11

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> AF055033

<400> 11

```

ggggaaaaga gctaggaaaag agctgcaaag cagtgtgggc tttttccctt tttttgctcc 60
ttttcattac cctcctccg ttttcacct tctccggact tgcgtagaa cctgcgaatt 120
tcgaagagga ggtggcaaag tgggagaaaa gaggtgttag ggtttggggg ttttttgttt 180
ttgtttttgt tttttaattt cttagatttc acattttctc ccacctctc ggctgcagcc 240
aacgcctctt acctgttctg cggcgccgcg caccgctggc agctgagggt tagaaagcgg 300
ggtgtatttt agattttaag caaaaaattt aaagataaat ccatttttct cteccacccc 360
caacgccatc tccactgcat ccgatctcat tatttcgggtg gttgcttggg ggtgaacaat 420
tttggtgctt tttttccct ataattctga cccgctcagg cttagagggt tctccggcct 480
ccgctcactg cgtgcacct gcgctgccct gcttccccca acctgttgca aggttttaat 540
tcttgcaact gggacctgct cgcaggcacc ccagccctcc acctctctct acatttttgc 600
aagtgtctgg gggagggcac ctgctctacc tgccagaaat tttaaaacaa aaacaaaaac 660
aaaaaaaaat ccggggggccc tcttggcccc tttatccctg cactctcgct cctctgccc 720
accccgaggt aaagggggcg actaagagaa gatggtgttg ctaccgcgg tccctctgct 780
gctggccgcc tatgcggggc cggcccagag cctgggctcc ttcgtgcact gcgagccctg 840
cgacgagaaa gccctctcca tgtgcccccc cagccccctg ggctgcgagc tgggtcaagga 900
gccgggctgc ggctgctgca tgacctgcgc cctggccgag gggcagtcgt gcggcgtcta 960
caccgagcgc tgcgcccagg ggctgcgctg cctcccccg caggacgagg agaagccgct 1020
gcacgccctg ctgcacggcc gcgggggttt cctcaacgaa aagagctacc gcgagcaagt 1080
caagatcgag agagactccc gtgagcacga ggagcccacc acctctgaga tggccgagga 1140

```



```

gacctactcc cccaagatct tccggcccaa acacacccgc atctccgagc tgaaggctga 1200
agcagtgaag aaggaccgca gaaagaagct gacccagtc aagtttgtcg ggggagccga 1260
gaacactgcc ccccccgga tcatctctgc acctgagatg agacaggagt ctgagcaggg 1320
cccctgccgc agacacatgg aggcttccct gcaggagctc aaagccagcc cacgcatggt 1380
gccccgtgct gtgtacctgc ccaattgtga ccgcaaagga ttctacaaga gaaagcagtg 1440
caaaccttcc cgtggccgca agcgtggcat ctgctggtgc gtggacaagt acgggatgaa 1500
gctgccaggc atggagtacg ttgacgggga ctttcagtgc cacaccttcg acagcagcaa 1560
cgttgagtga tgcgtcccc cccaaccttt ccctcaccoc ctcccacccc cagccccgac 1620
tccagccagc gcctccctcc accccaggac gccactcatt tcatctcatt taagggaaaa 1680
atatatatct atctatttga ggaaaaaaaa aaaaaaaaaa aa 1722

```

<210> 12

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF055033

<400> 12

```

tccaccccag gacgccactc atttcatctc atttaaggga aaaatatata tctatctatt 60

```

<210> 13

<211> 1411

<212> DNA

<213> Homo sapiens

<300>

<308> AK001166

<400> 13

```

aaacaaaagag atgccacccc tgtgtgatgg ctttgggtacc cgaacactga tgggttcagac 60
attttcccggt tgcattcttgt gttccaagga tgaagtggac ttggatgagt tattagctgc 120
tagattggta acgtttctga tggacaatta ccaggaaatt ctgaaagtcc ctttggcctt 180
gcagacctct atagaggagc gtgtggctca tctacgaaga gtccagataa aatacccagg 240
agctgatatg gatatactt tatctgctcc atcattttgc cgtcaaatta gtccagagga 300
atltgaatat caaagatcat atggctctca ggaacctctg gcagccttgt tggaggaagt 360
cataacagat gccaaactct ccaacaaaga gaaaaagaag aaactgaagc agtttcagaa 420
atcctatcct gaagtctatc aagaacgatt tectacacca gaaagtgcag cacttctgtt 480
tcctgaaaaa cccaaaccga aaccacagct gctaattgtgg gcaactaaaga agcctttcca 540
accatttcaa agaactagaa gttttcgaat gtaataatac ttccacagca acaggtgcta 600
gagaccactg ttgttgtttt gagtgaatgg tgggttaggag aaagactttg gtggtggaag 660
aaagaaaagc ataaaaacaaa gactactgaa atatagataa agattgcctt agttttttaa 720
aatgtttggc cattagtatt tttataaaac tcaatgctag ttttaagtgt ataaattggg 780
taaaatttat gagtcaaata tatagtata atgttaacat gtttgtaatt gctacagaat 840
ttaagggtat ttttatctct gtgctttctt tttcatgggtg tttattaaat aattgtgtat 900
atacatccta gctactgata tctttattat agccttaaga cttaatttta agtcttaaaa 960
atagcgtgta tacttgaata agaaagacac tgggtactgt tactgtgatg ctattgactt 1020
agtagccaat tatcatttct cctgtataaa ttccagtttt tattgctgca cataaatttt 1080
ttaatgtctt atattgtgat agctatgtct tttattgcag atttattgga tgttatgaca 1140
gattttacta aagctagtgt ttttataaca tatatattag ttgatgttta cctataagtg 1200
gagtagatth tcatctgcct gcaatgggat aatttcagtc ttagctaaaa atggaaagtt 1260

gaactggata aattcttttg gtacccttag acctctgatt ctaagtcaaa tgcaaatggg 1320
ttaaataaaa tgagactact tcctttataa atatatattt atccttttga aagtaagtga 1380
aatgtaaata aacttatttt ttttaaaaat g 1411

```

<210> 14

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AK001166

<400> 14

acccttagac ctctgattct aagtcaaag caaatgggtt aaataaaatg agactacttc 60

<210> 15

<211> 2352

<212> DNA

<213> Homo sapiens

<300>

<308> AL049367

<400> 15

```

ggcaaacccc ttttaaaatc taatgtctgg gctttgagta ttagctcatt taggggtggac 60
aaatgcatta ctgttttcaa actgttcaca tttattcagt atttctccaa gttgctatct 120
actcagcctt atgaatgcc ctcgttttcc taaggccatg tgaaaatcac ggcactgccc 180
ttagccttgt gtcactctgt ttttctgtct gcgatatgcc cagttcccaa atcaattata 240
ggtacctgtt taggagagag gaagatttta cctctcaaag ggtgagattt gaaatttaca 300
ctaaaaagac aactttacat ttaatgcttc acttaatgag acattctttt ttttataagt 360
ctatttttct actcagtttc agaacactaa tctgattttc actctgattt ttaacgtttc 420
tttaaatatt tataatgtag cttctttcaa aatattttca tgaaaaatta cttttattat 480
accattatgt gcatgttatt ggtagcaggc atagtttatt atttagtact gaaacatgct 540
cttttaccta acagtaaaca agtatgtttt gatatatatc tgttaatatg cttatagtgg 600
taagaaatgg acttgagggt ccaggagatt tcatttttatt caccctgggt agatacaata 660
aaggctatga gtataaatac ataacttcct aaccagggtg agggcatggt catgaatata 720
aaatcttttg atgctggacc caagagagga aaagtgttag ctaaatgttg atttacttat 780
aactagacgt ctatgtgaga aaatatatgt atacatatat atgatatgca gaagtcactt 840
tttttatcag gctttattct ccttacaaag ccacagttta actgtctgca acagttgggt 900
tatgttaatg atagacaaat acccagtggt tgttactttt tccaactacc actgtaatga 960
taatctttct cacgtatata catgcaactt cttggcttca tttccatgaa gctgtttcaa 1020
tatattcagt atactttgtc cttaatgctg cttctgttaa cagtgatctc tttctttttt 1080
tcattcttat atcttcatta gttcatcata aatctgtcca gttgaggcct caggaccacg 1140
gcatgatttc atgactccga agtattttac agaaacattt tttaaataag ggaaatatat 1200
tatataccag atggttcaca agtgatggct catagctagt tttttttttt tcttctaaaa 1260
aatgtcaggt ttttaaaatc atttacctta ttaaaatgaa aagtgccata cttaactttt 1320
aaaggaaaga cctgacttgc tttttctcta ttttagactgt ttttgtactt tactaatctt 1380
taaactatca ggaaaaaac caaaacttta taccaatgat ttagtaattt tgaggcatag 1440
ggtagcttac gtagtggagg atgtgccaaa tattctcttc aaatgccacc ttctcaattt 1500
ataactaaaa tagtggtatc tgactaattc ctctgaattt tgatgtaaga tctatatagg 1560
cccccaaat gatcgtagta catgccagtc atttctcagt gaaataaata caataccaga 1620
gtacattatg gggtttattg ctttctttta tggtagacct gttaatgggg aaaaaatata 1680
tcaaatcaaa tagaatctta tatctgtatg ttaaaataga gcacttacct gaagtcagt 1740
gcctggatca tagccctgga tcatttccca gtctgtcttg tgctgtgtga ccttgacaaa 1800
ggcgcttcat ctctctgggc ctctatttct ccatttgtaa aacaagtggc tgcagtagat 1860
gatggctgag agcccttcct gttcccagat gccttgggtc aaagacccca cccctctgct 1920
ggtcctgcca acgtgttggg gctataagct gcttcagata taaaattggg ttatctataa 1980
tggtttgttca ttttaatagct tctaaaaggc ctttttgtaa tacagtgtt tttttctagt 2040
tttatggact tgattactgt aataatgtct tgttttttagc catgtaacta caaacagata 2100
ttctcttgat gtcttagtaa atttgcattt gatatatcat tgatgagatt ttgttgttat 2160
gtaatatctt ttggctacgc atctgtccag catcttatta accataatac tgtgatcatt 2220
atgttgaaat atgtcctatg gaaagaataa aagcatgtac ttcacagcta gcatgttcac 2280
agatttgaaa gaagtttcat taaaagcacc attgctttct gtaaaaaaaaa aaaaaaaaaa 2340
aaaaaaaaaa aa 2352

```

<210> 16

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AL049367

<400> 16

atttggaat atgtcctatg gaaagaataa aagcatgtac ttcacagcta gcatgttcac 60

<210> 17

<211> 1130

<212> DNA

<213> Homo sapiens

<300>

<308> AL080235

<400> 17

```

ggtcgccgca ccggccgcct ccggcccgcc gccgccccca gccgcccgc ccgccaccgcc 60
ggggcgccca ccgcgctgcc agcctacccc gccggccgagc cgcggggggcc gctgtggctg 120
cagggcgagc cgctgcattt ctgctgccta gacttcagcc tggaggagct gcagggcgag 180
ccgggctggc ggctgaaccg taagcccatt gagtccagc tgggtggcctg cttcatgacc 240
ctggctcatc tgggtgtggag cgtggccgcc ctcatctggc cggtgcccat catcgccggc 300
ttcctgcca accgcatgga acagcgccgg accaccgcca gcaccaccgc agccaccccc 360
gccgcagtgc ccgcagggac caccgcagcc gccgcgcgc ccgcgctgc cgcggccgcc 420
gcggccgcta cttcgggggt ggcgaccaag tgaccgctc cgtcctccc tgtgtccgtc 480
ctgtgtccgc gcgcgcgggt gcctttcccg ccggggactc ggccgggtgtg cttcgtgctg 540
tagttatcgt tagttcctct tcccgagatg gggccgcca gagaccagc cgcctttgaa 600
aagcaagggt tgtgctgcgc ttccagttcc gaaaagcaga tgtttaagcc cttggactga 660
gggtgggata gcagctccga agacggagag gagggaaatg gggecccttc cctctattg 720
catccccctg ccgactcct tccccgcacc cactgcccct agattcatgg cagaaaatga 780
ccaaatcctg tgtatttgtt ttatatattt aataactgtt ttaaatgaaa gtttttagtaa 840
aaaaaataca aaacaaaaag attaaattgc tattgtgtga gtaagagaag ctctttgtat 900
ctgaacatag ttgtatttga aatttgtggg tttttaattt atttaaaatt ggggggaggg 960
catgggaagg atttaacacc gatataattg taccgctgaa aatgaacttt atgaaccttt 1020
tccaagttga tctatccagt gacgtggcct ggtgggcgtt tcttcttgta cttatgtggg 1080
tttttggtt ttaatacaga cattttcctc caaaaaaaaa aaaaaaagg 1130

```

<210> 18

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AL080235

<400> 18

ctttgaaaag caaggtttgt gctgcgcttc cagttccgaa aagcagatgt ttaagccctt 60

<210> 19

<211> 2498

<212> DNA

<213> Homo sapiens

<300>

<308> AL137540

<400> 19

```

gctgaaacga cagtcttgtc cctgtcagag aaatgacctg aacgaagagc ctcaacattt 60
tacacactat gcaatctatg atttcattgt caagggcagc tgcttctgca atggccacgc 120
tgatcaatgc atacctgttc atggcttcag acctgtcaag gcccaggaa cattccacat 180
ggtccatggg aagtgtatgt gtaagcaca cacagcagc agccactgcc agcactgtgc 240
cccgttatac aatgaccggc catgggaggg agctgatggc aaaacggggg ctccaacga 300

```

```

gtgcagaacc tgcaagtgt atgggcatgc tgatacctgt cacttcgacg ttaatgtgtg 360
ggaggcatca gggaatcgta gtggtggtgt ctgtgatgac tgtcagcaca acacagaagg 420
acagtattgc cagagggtgca agccaggctt ctatcgtgac ctgcccagac ccttctcagc 480
tccagatgct tgcaaacgt gttcctgcca tccagtagga tcagctgtcc ttcttgccaa 540
ctcagtgacc ttctgcgacc ccagcaatgg tgactgccct tgcaagcctg ggggtggcagg 600
gcgacgttgt gacagggtgca tgggtgggata ctggggcttc ggagactatg gctgtcgacc 660
atgtgactgt gcggggagct gtgaccctat caccggagac tgcatacagca gccacacaga 720
catagactgg tatcatgaag ttctgactt ccgtcccgtg cacaataaga gcgaaccagc 780
ctgggagtgg gaggatgctc aggggttttc tgcacttcta cactcaggta aatgcgaatg 840
taaggaacag acattaggaa atgccaaggc attctgtgga atgaaatatt catatgtgct 900
aaaaataaag attttatcag ctcatgataa aggtactcat gttgaggta atgtgaagat 960
taaaaaggtc ttaaaatcta ccaaaactga gattttccga ggaaagcga cattatatcc 1020
agaatcatgg acggacagag gatgcacttg tccaatcctc aatcctggtt tggaatacct 1080
tgtagcagga catgaggata taagaacagg caaactaatt gtgaatatga aaagctttgt 1140
ccagcactgg aaaccttctc ttggaagaaa agtcatggat attttaaaaa gagagtgcac 1200
gtagcattaa gatggatagc acataatggc acttgtctat gtacaaaaca caaactttag 1260
agcaagaaga cctcagacag gaaactggaa ttttttaaag tgccaaaaca tatagaaatg 1320
tttgaatgca tgggtottat ctaacttata tcttctggac ccatgtttta atacagtttt 1380
atttcatgaa gagaaatgaa aacccttaca ctgatatctg ttttctatgg gactgattct 1440
gaaattctta actattaaga atattttaat agcagcatga catttagcag taatccatta 1500
agggcagtac ctctaacaag gacgccttcc agcttcagcg atgttactta cgtttgatgc 1560
tacttaaagt aatgaatgac gttttaagga atccctaacc ctactatcag aaaagggtgt 1620
tgtaaagag ccttctcttg tgtgttacgc atgaactttg gtctgtagg gttaaattgga 1680
acctctccat gtgtatatag tatttccttg tataaagcac ttactacct accacttgtg 1740
ttgtgaacgt ttggtgactg ctgttgaaag aaggaaaagg gtgtgtgaga aagcctactg 1800
aagcagcagc actgccacta catgtggaca aaagtgacca tataaaagaa gttgtgctat 1860
ttaactctga atacttggag aaactagggt aagatgcaac cagaaaggag aatatgtatg 1920
cgtgaagtct cagctttgag ctggaggcta gattccaaga tgacagccat gatgaaactt 1980
tttaaaaaac taaaccagaa gagacttta aataagagaa agaaatcata aatgtagaca 2040
tatgcttggc taaaggggaa atggacttta aattttaaag agctcatttg caatgcactt 2100
gtatacactt caaaaattat tgtagacaca gaatttgta tatttttgtg cttagtattt 2160
aaacctgaac attgaaacag ttttccctct tgtctttctt aacagtaata gtcattatat 2220
ttacctgttt ttttaacaca tgtatgtgat agtcaaaaaa tcacagtttt tcattattat 2280
tcacttcttg taccacagca taaccactat acatagtttc ttttgtactt gaatatacaa 2340
aacatgaaca cagtgcata tgaataattt cacatacaga accttttttt ctctgaagtc 2400
ctgtggactt gcaaatatat atatatattg ctttgttaat ttgtttttat atttcatata 2460
tgtaataaag gaatatgatc tgaaaaaaa aaaaaaaa 2498

```

<210> 20
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL137540

<400> 20
 tggaggctag attccaagat gacagccatg atgaaacttt ttaaaaaact aaaccagaag 60

<210> 21
 <211> 914
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL160131

<400> 21
 cgcaccgcag gagcaacggt tggctcctgcg gctgtgatgt cgggtgttgag gccctggac 60
 aagctgcccg gcctgaacac ggccaccatc ttgctggttg gcacggagga tgcctctctg 120
 cagcagctgg cggactcgat gctcaaagag gactgcgcct ccgagctgaa ggtccacttg 180

```

gcaaagtccc tccctttgcc ctccagtgtg aatcggtccc gaattgacct gatcgtgttt 240
gtggttaatac ttcacagcaa atacagtctc cagaacacag aggagtcctt gcgccatgtg 300
gatgccagct tcttcttggg gaaggtgtgt ttctctcgca caggtgctgg gcgggagagc 360
cactgcagca ttcaccggca caccgtggtg aagctggccc acacctatca aagccccctg 420
ctctactgtg acctggaggt ggaaggcttt agggccacca tggcgagcgc cctggtgctg 480
gtgctgcaga tctgtgctgg ccacgtgccc ggtgtctcag ctctgaacct gctgtccctg 540
ctgagaagct ctgagggccc ctccctggag gaacctgtgag ggtggctggc cctgggctg 600
cccccttctca tggcttctgt ctgactccat aaacattctc tgttgaggat gtccagtcag 660
ggcttgacag gcccaggctc agcccggcgt ggctgggaag gttccctgca gtgccagtgc 720
tgcagcaggg agagctgggc agaagcagcg agggggccca gctggcgaga ctgtagcccc 780
ctcccactcc cacactcact cttgcagagc ctgtgtcttt aagcagctgg cgtgttacat 840
ctccatttaa ggtttccttt gaacaaaagg tctgtggcta aaaaaagttt aaaaatcact 900
ggtctcatcc acca 914

```

<210> 22

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AL160131

<400> 22

agctggcgtg ttacatctcc atttaagggt tcctttgaac aaaaggtctg tggctaaaaa 60

<210> 23

<211> 4753

<212> DNA

<213> Homo sapiens

<300>

<308> D13642

<400> 23

```

cttcaatcaa gtagccttcc cactgcagta cacaccagg aaatttgtca tccaccctga 60
gagtaacaac cttattatca ttgaaacgga ccacaatgcc tacactgagg ccacgaaagc 120
tcagagaaag cagcagatgg cagaggaaat ggtggaagca gcaggggagg atgagcggga 180
gctggccgca gagatggcag cagcattcct caatgaaaac ctccctgaat ccattcttgg 240
agctcccaag gctggcaatg ggcagtgggc ctctgtgatc cgagtgatga atcccattca 300
agggaacaca ctggaccttg tccagctgga acagaatgag gcagctttta gtgtggctgt 360
gtgcagggttt tccaacactg gtgaagactg gtatgtgctg gtgggtgtgg ccaaggacct 420
gatactaaac ccccgatctg tggcaggggg ctctgtctat acttacaagc ttgtgaacaa 480
tggggaaaaa ctggagtttt tgcacaagac tcctgtggaa gaggtccctg ctgctattgc 540
ccattccag gggaggggtg tgattggtgt ggggaagctg ttgcgtgtct atgacctggg 600
aaagaagaag ttactccgaa aatgtgagaa taagcatatt gccattata tctctgggat 660
ccagactatt ggacataggg taattgtatc tgatgtccaa gaaagtttca tctgggttcg 720
ctacaagcgt aatgaaaacc agcttatcat ctttgtgat gatacctacc cccgatgggt 780
cactacagcc agcctcctgg actatgacac tgtggctggg gcagacaagt ttggcaacat 840
atgtgtgggtg aggtccccc ctaacaccaa tgatgaagta gatgaggatc ctacaggaaa 900
caaagccctg tgggaccgtg gcttgcctca tggggcctcc cagaaggcag aggtgatcat 960
gaactacat gtgcggggaga cggtgctgtc cttgcagaag accacgctga tccttgagg 1020
ctcagaatca cttgtctata ccaccttgtc tggaggaatt ggcacctctg tgccattcac 1080
gtcccatgag gaccatgact tcttccagca tgtggaaatg cacctgcggt ctgaacatcc 1140
ccctctctgt gggcgggacc acctcagctt tcgtctctac tacttccctg tgaagaatgt 1200
gattgatgga gacctctgtg agcagttcaa ttccatggaa cccaacaaac aaaagaacgt 1260
ctctgaagaa ctggaccgaa ccccaccgaa agtgtccaag aaactcgagg atatccggac 1320
ccgctacgcc ttctgagccc tcctttcccg gtggggcttg ccagagactg tgtgttttgt 1380
ttccccacc accatcactg ccacctggct tctgccatgt ggcaggaggg tgactggata 1440
attaagactg cattatgaaa gtcaacagct ctttccctc agctcttctc ctggaatgac 1500
tggcttcccc tcaaattggc actgagattt gctacacttc tccccacctg gtacatgata 1560
catgacccca ggttcagtg tagaacctga gtccccatt ccccaaagcc atccctgcat 1620

```

tgatatgtct	tgactctcct	gtctactttt	gcacacaccc	ttaattttta	attggttttc	1680
ttgtaaatac	agttttgtac	aatgtttatct	ctgtgtggagg	aaggaggcag	gctgtgtgtg	1740
gactgggtag	ggtatagtat	cactcctgag	ttccactgct	ctagaatcta	accagaaata	1800
gaaacctagt	ttttaagggtg	actggcatcc	atgtgtcttg	ttctggagat	gaggatgtag	1860
gtgggagggt	tgaacccaag	ttagagcagg	aagaactgag	tagactcctt	ccttccagat	1920
accgacttgg	acttgcggca	ctctgtggct	ccccaccccc	aggtctgtgg	tggtttcttt	1980
gttttttctc	ggttcttttt	gctgtgctga	tgaacatga	cctcaataac	catgtgtata	2040
cccacccctc	ttcccactgg	gtattgagga	aggggtggctg	attcttctct	ctcttctact	2100
ctgaggatgt	tagtatgggg	attttagcat	gaattccagc	tggggagtct	taacagatgc	2160
cccttttact	gatagagcac	ctaaagcgat	ctttggctcc	ataggaccat	aggaagggtc	2220
agtcacagaag	aacctagata	ctgccctgcc	cctgagaact	gtgtatatgt	ggggcctgtc	2280
tgcagcacc	atctcagggtg	ggttccagag	ggccttttagg	gtataatgag	agcctgttag	2340
gtggaagagg	cccagttcca	gaaatgttcc	agccaccccc	tgagaattcc	tcctgttttag	2400
ttgtgtggga	agccctcgtc	ttccaggctg	tccttgccgc	ttgaacctgg	agaagtgagc	2460
tcactgttct	caatacttca	caaatgtaaa	actttctttc	gtctgcatgt	gctcagccat	2520
ctaaattgag	caaatgatct	ggtgagcact	gggttagaat	caggaatggg	ggaatacaat	2580
ctgaacctct	cagagcccag	aacagagggg	tcctgacact	gtgacactgt	ctcctggaac	2640
taagtatctc	ttgaatcatg	acttgggttt	agatcagtc	agagagaccc	aggttttgcc	2700
aggaatcgaa	tccttaata	acatgttttt	ttctcactta	gctcatgaat	ttgcatagta	2760
gacagtagtt	ctgaattaga	ttttgaaaac	ctaatttcag	ggctcatttt	ttcctgtggc	2820
cctaaatcca	ttctatcaaa	ttgtgtgata	ctgacatgca	gtcatctgag	gaactcagcg	2880
tagatacttg	agcagctcct	cgccctcttt	ctaactcaag	tttgactaaa	atacatacac	2940
tccgtacaga	aggtaggggg	ttatgtaaga	aaggaaaacc	taatctatgg	aatcaggagt	3000
tgtcaccacc	gagcttcctc	tgggaagtctg	cccatcagct	tgcttgttct	ctgttaagag	3060
gaagggctag	gacaaggatt	tgggcttgaa	tatgtggaaa	ggaattttca	tagttgttgc	3120
tgcaggacct	acaaaagtgt	aaaattagat	tggatgtgac	tcaatgacaa	gtcccatctg	3180
tgtaatgtgt	aaggggacct	gattgactcc	tgtgggttga	ttgagcaacc	aggtaaatag	3240
agacctctct	ccagcttttg	caaaacccat	cagaggctgc	tgcagaactc	agacagaggg	3300
atctgcccct	gggtttgtct	ccatcctggt	ccattgctaa	gcccttgtga	cttggatcct	3360
aggactgaaa	agttttttagc	tgcctcagct	ttcccctgac	cttactggca	gagggtctgc	3420
agatgtttcc	tttgggaagat	ctcttgccaa	gaatagcatt	cctttggagg	aggggggttc	3480
tagttggaat	gttgcttttc	ttggtttagtg	taaatgtatt	gctagtgaga	cagctgccgg	3540
cgctggaaaa	ggctcgtctc	acaggagagag	tgctgggtccc	cagaatgtgt	gctgttccca	3600
cgctgctgcc	tttcttgagc	ttgttagagg	aaagccagaa	aggcattcag	atgggatcag	3660
tctggctttc	aaattttttt	taattcctaa	gttctgtttt	attttttaat	tttttaaaaa	3720
aaattttatt	agagacagtc	tctctctctt	gcctagctgg	gagtgcagtg	gagtgatcat	3780
agctcactga	ggcttgaaact	cctgggctcg	agcaatccac	ctcagcctcc	agagtagggg	3840
agactacaga	tgtgtgccac	catactcagc	tagttttttaa	actttcgtag	agacaggggtc	3900
tcctgtgtgt	gccagggctg	gcctcgaact	cctgacctca	aaaaatcttc	ctgccttggc	3960
ctcccagcgc	tttgagaggc	tgaggcagga	ggatcccttg	agcccaggag	tttgagacca	4020
gcctgggcaa	catgacaaaa	ccccatctct	ccaaaaatac	aaaaattggc	caggcatggg	4080
ggtgcacact	tgtagtccca	gtaattaggg	ggctgagaca	ggaggatcac	ttcagcctat	4140
gagtttgagg	ctgcagttag	ctgtgattgc	gccactacac	tcagcctggg	atgacaggac	4200
gaaacctgtc	tcaaaaacac	caaaaaacaa	aaaccgggtct	cctgggggtca	tggtagcaca	4260
aacgcacatg	actgagtgtc	caggggttct	gaggcttgtc	cgctgacctg	gggctctggc	4320
cctgggagat	ctgggggacc	tgctgtccta	tatgtgatgc	tttgaaagaa	aggggcatca	4380
ttccaagcca	agaggcccca	gagagggcac	cgtaggggtgt	tcaggcttct	gtgaggcccc	4440
agtgagatcc	tgtggctgtg	cccccatcac	ctccaccac	tctgcctcc	cactagctgc	4500
ccaacggatg	aatcaacgcc	ttggcagagt	tttccagcag	ggccttgacg	agagtgtgtg	4560
tgacctgtgt	ggccactgcc	ttggggacgg	gtgaggagtt	agcctggaac	attccagcgt	4620
gggcattatt	gtcctgttgc	aagttcaggg	caaaaccagg	aatccagttt	tgtcgatcca	4680
attgagaaaa	catttcatga	acaactactt	gtggcatgca	ttggcactcg	gaataaagcg	4740
cactattgtc	act	4753				

<210> 24

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> D13642

<400> 24
aaaccaggaa tccagttttg tccatccaat tgagaaaaca tttcatgaac aactacttgt 60

<210> 25
<211> 2591
<212> DNA
<213> Homo sapiens

<300>
<308> D25328

<400> 25
cccgagcgtg cggctcccct cggcctctc gccatggacg cggacgactc cggggccccc 60
aagggctcct tgcggaagtt cctggagcac ctctccgggg cgggcaaggc catcggcgtg 120
ctgaccagcg gcgggggatgc tcaaggatat aacgctgccg tccgtgccgt ggtgcgcag 180
ggtatctacg tgggggcca ggtgtacttc atctacgagg gctaccaggg catgggtggac 240
ggaggctcaa acatcgaga ggcgactgg gagagtgtct ccagcatcct gcaagtgggc 300
gggacgatca ttggcagtgc gcggtgccag gccttccgca cgcgggaagg ccgcctgaag 360
gctgcttgca acctgctgca gcgcggcatc accaacctgt gtgtgatcgg cggggacggg 420
agcctcaccg gggccaacct ctcccggaag gagtggagtg ggctgctgga ggagctggcc 480
aggaacggcc agatcgataa ggaggccgtg cagaagtacg cctacctcaa cgtggtgggc 540
atggtgggct ccacgcacaa tgatttctgc ggcaccgaca tgaccatcgg caccgactcc 600
gccctgcaca ggcacatcga ggtcgtcgac gccatcatga ccacggccca gagccaccag 660
aggaccttcg ttctggaggt gatgggacga cactgtgggt acctggccct ggtgagtgc 720
ttggcctgcg gtgcggactg ggtgttctct ccagaatctc caccagagga aggtgggag 780
gagcagatgt gtgtcaaaact ctcgagaaac cgtgcccgga aaaaaaggct gaattattat 840
attgtggctg aaggagcaat tgatacccaa aataaaccca tcacctctga gaaaatcaaa 900
gagcttgtcg tcacgcagct gggctatgac acacgtgtga ccacctcgg gcacgtgcag 960
agaggaggga ccccttcggc attcgacagg atcttggcca gccgcaggg agtggaggca 1020
gtcatcgcc tgcagaggg caccocggac acccagctt gcgtcgtgtc actgaacggg 1080
aaccacgccc tgcgcctgcc gctgatggag tgcgtgcaga tgactcagga tgtgcagaag 1140
gcgatggacg agaggagatt tcaagatgcg gttcgactcc gagggaggag ctttgcgggc 1200
aacctgaaca cctacaagcg acttgccatc aagctgccgg atgatcagat ccaaagacc 1260
aattgcaacg tagctgtcat caacgtgggg gcaccgcgg ctgggatgaa cgcggccgta 1320

cgctcagctg tgcgcgtggg cattgcgcac ggccacagga tgctcggcat ctatgatggc 1380
tttgacggct tgcgcaaggg ccagatcaaa gaaatcggct ggacagatgt cgggggctgg 1440
accggccaag gaggtccat tcttgggaca aaacgcgttc tcccggggaa gtacttggaa 1500
gagatcgcca cacagatgcg caccgcacagc atcaacgcgc tgctgatcat cgggtggattc 1560
gaggcctacc tgggactcct ggagctgtca gccgcccggg agaagcacga ggagtctgt 1620
gtccccatgg tcatggttcc cgtactgtg tccaacaatg tgccgggttc cgatttcagc 1680
atcggggcag acaccgcct gaacactatc accgacacct gcgaccgcat caagcagtc 1740
gccagcggaa ccaagcggcg cgtgttcatc atcgagacca tggggcgcta ctgtggctac 1800
ctggccaaca tgggggggct cgcggccgga gctgatgccg catacatttt cgaagagccc 1860
ttcgacatca gggatctgca gtccaacgtg gagcacctga cggagaaaaat gaagaccacc 1920
atccagagag gccttgtgct cagaaatgag agctgcagtg aaaactacac caccgacttc 1980
atttaccagc tgtattcaga agagggcaaa ggcgtgtttg actgcaggaa gaacgtgctg 2040
ggtcacatgc agcagggtgg ggcaccctct ccatttgata gaaactttgg aacaaaaatc 2100
tctgccagag ctatggagtg gatcactgca aaactcaagg agggccgggg cagaggaaaa 2160
aaatttacca ccgatgattc catttgtgtg ctgggaataa gcaaaagaaa cgttattttt 2220
caacctgtgg cagagctgaa gaagcaaacg gattttgagc acaggattcc caaagaacag 2280
tggtggctca agctacggcc cctcatgaaa atcctggcca agtacaaggc cagctatgac 2340
gtgtcggact caggccagct ggaacatgtg cagccctgga gtgtctgacc cagtcccgc 2400
tgcattgtgc tgcagccacc gtggactgtc tgtttttgta acacttaagt tattttatca 2460
gcactttatg cagctattat tgacattaat acctaatcgg cgagtgcaca tctgccccac 2520
cagctccagt gcgtgctgtc tgtggagtgt gtctcatgct ttcagatgtg catatgagca 2580
gaattaatta a 2591

<210> 26
<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> D25328

<400> 26

tattttatca gcactttatg cacgtattat tgacattaat acctaatacgg cgagtgccca 60

<210> 27

<211> 2573

<212> DNA

<213> Homo sapiens

<300>

<308> D50402

<400> 27

```

gaatcggccg atgtgaaccg aatgttgatg taagaggcag ggcactcggc tgcggatggg 60
taacagggcg tgggctggca cacttacttg caccagtgcc cagagagggg gtgcaggctg 120
aggagctgcc cagagcaccg ctacactcc cagagtacct gaagtcggca tttcaatgac 180
aggtgacaag ggtccccaaa ggctaagcgg gtccagctat ggttccatct ccagcccgac 240
cagcccgacc agcccagggc cacggcaagc acctccaga gagacctacc tgagtgagaa 300
gatccccatc ccagacacaa aaccgggcac cttcagcctg cggaagctat gggccttcac 360
ggggcctggc ttctcatga gcattgcttt cctggacca ggaaacatcg agtcagatct 420
tcaggctggc gccgtggcgg gattcaaact tctctgggtg ctgctctggg ccaccgtgtt 480
gggcttgctc tgccagcgac tggtgcacg tctgggcgtg gtgacaggca aggacttggg 540
cgaggctctg catctctact accctaaggt gccccgcacc gtctctctgg tgaccatcga 600
gctagccatt gtgggctccg acatgcagga agtcatcggc acggccattg cattcaatct 660
gctctcagct ggacgaatcc cactctgggg tggcgctctc atcaccatcg tggacacctt 720
cttcttctct ttctcagata actacgggct gcggaagctg gaagcttttt ttggactcct 780
tataaccatt atggccttga cctttggcta tgagtatgtg gtggcgcgct ctgagcaggg 840
agcgcttctt cggggcctgt tctgcccctc gtgcccgggc tggggccacc ccgagctgct 900
gcaggcgggtg ggcattgttg gcgccatcat catgcccac aacatctacc tgcactcgge 960
cctggtaacg tctcgagaga tagaccgggc ccgocgagcg gacatcagag aagccaacat 1020
gtacttcctg attgaggcca ccatcgccct gtcogtctcc tttatcatca acctctttgt 1080
catggctgtc tttgggcagg ccttctacca gaaaaccaac caggctgogt tcaacatctg 1140
tgccaacagc agcctccacg actacgcaa gatcttcccc atgaacaacg ccaccgtggc 1200
cgtggacatt taccaggggg gcgtgatcct gggctgcctg ttcggecccc cggccctcta 1260
catctggggc ataggtctcc tggcggtgg gcagagctcc accatgacgg gcacctacgc 1320
gggacagttc gtgatggagg gcttcctgag gctgcgtggg tcacgcttcg ccctgtcct 1380
cctcaccgcg tctgcgcca tctgcccac cgtgctcgtg gctgtcttcc gggacctgag 1440
ggacttgctg ggctcaatg atctgctcaa cgtgctgcag agcctgctgc tccggttcgc 1500
cgtgctgccc atcctcacgt tcaccagcat gccaccctc atgcaggagt ttgccaatgg 1560
cctgctgaac aaggctcgtc cctcttccat catggtgcta gtctgcgcca tcaacctcta 1620
cttcgtggtc agctatctgc ccagcctgcc ccaccctgcc tacttcgggc ttgcagcctt 1680
gctggccgca gcctacctgg gcctcagcac ctacctgtc tggacctgtt gccttgcca 1740
cggagccacc tttctggccc acagctccca ccaccacttc ctgtatgggc tcttgaaga 1800
ggaccagaaa ggggagacct ctggctaggc ccacaccagg gcctggctgg gagtggcatg 1860
tatgacgtga ctggcctgct ggatgtggag ggggcgctg caggcagcag gatggagtgg 1920
gacagttcct gagaccagcc aacctggggg ctttagggac ctgctgtttc ctagcgcagc 1980
catgtgatta ccctctgggt ctcagtgctc tcatctgtaa aatggagacg ccaccaccct 2040
tgccatggag gttaagcact ttaacacagt gtctggcact tgggacaaaa acaacaaaac 2100
aaacaaaaaa catttcaaaa ggtatttatt gagcacctgc aggcgtgacc tgacagccca 2160
aggggtgggtg ggggtgagggc ttgaggactt gggcgggaca caggctccaa actggagctt 2220
gaaatagtgt ctgatgaatg ttaaattatc tatctatcta tttatttatt tttttgagac 2280
agggaaaggg tctccctctg ttgccaaggg tggagtgcag tggcgcaatc ttaactcatt 2340
gcaacctcca ccttctgggt tcaagcgatt ctctttatct agccccggga gtggcgcgcg 2400
ccaccacgcc cagctaattt gtgtattttc agcagagacg ggggttgcca tgctggccag 2460
gctgggtctcg aactgctgga ttcaagtgat ccgcccctct ccgtctccca aagtgctggg 2520

```


aattacaggc gtgagccacc aaaacccggc ctgattaaag ttaaataaat acg 2573

<210> 28

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> D50402

<400> 28

tggagggttaa gcactttaac acagtgtctg gcacttggga caaaaacaaa caaacaaaca 60

<210> 29

<211> 3672

<212> DNA

<213> Homo sapiens

<300>

<308> L27560

<400> 29

acatgtgcat	atttcattcc	ccaggcagac	atTTTTtaga	aatcaatata	tgccccaata	60
ttggaaagac	ttgttcttcc	acggtgacta	cagtacatgc	tgaagcgtgc	cgtttcagcc	120
ctcattttaat	tcaatttgta	agtagcgac	gagcctctgt	gggggaggat	aggctgaaaa	180
aaaaaagtgg	gctcgtatTT	atctacagga	ctccatatag	tcatatatag	gcataataat	240
ctatgctttt	tctttgtttt	tttctttctt	cctttctttc	aaaggtttgc	attaactttt	300
caaagtagtt	cctatagggg	cattgaggag	cttctctcatt	ctgggaaaac	tgagaaaacc	360
catattctcc	taatacaacc	cgtaatagca	tttttgctg	cctcgaggca	gagtttcccg	420
tgagcaataa	actcagcttt	tttgtggggc	acagtactgg	atTTgacagt	gattccccac	480
gtgtgttcat	ctgcacccac	cgagccaggc	agaggccagc	cctccgtggg	gcacacagca	540
cgcgcctcag	tccatcccat	tttagtcttt	aaaccctcag	gaagtcacag	tctccggaca	600
ccacaccaca	ttgagoccaa	cagggtccacg	atggatccac	ctagtcccac	cccagccttt	660
ttctttcatc	tgaacagaat	gtgcattttt	ggaagcctcc	ctcactctcc	atgctggcag	720
agcaggagggg	agactgaagt	aagagatggc	agagggagat	ggtggcaaaa	aggtttagat	780
gcaggagAAC	agtaagatgg	atggttccgg	ccagagtcca	tgtggggagg	aacagagggc	840
tgaagggaga	gggggcctgac	tgttccatcc	tagctttggc	acaaagcagc	agaaaggggg	900
aaaagccaat	agaaattttc	ttagcttccc	caccatatgt	atTTtcatgg	atTTgagagg	960
aaagagagga	aaatggggga	atgggttgca	aaatagaaat	gagcttaatc	caggccgcag	1020
agccagggaa	ggtgagtaac	cttaggaggg	tgctagactt	tagaagccag	ataggaagaa	1080
tcagtctaaa	ctggccatgc	tttggaaggg	acaagactat	gtgctccgct	gcccaccttc	1140
agcctgcaat	gagggactga	ggcccacgag	tctttccagc	tcttcccca	ttctggccag	1200
tccttgcatc	ctccctgggg	tggaggatgg	aaggaaagct	gggacaagca	gggaacgcac	1260
gattcaggga	tgctgtcact	cggcagccag	attccgaaac	tccattcttc	caatgacttc	1320
ctcaaccaat	gggtggcctt	gtgactgttc	tttaaggctg	aagatatcca	ggaaaggggg	1380
cttgacact	ggccaaggag	accccttcgt	gctgtggaca	cagctctctt	cactctttgc	1440
tcatggcatg	acacagcgga	gaccgcctcc	aacaacgaat	ttggggctac	gaagaggaat	1500
agcgaaaaag	caaactctgt	tcaactgatg	ggaaccctat	agctatagaa	cttgggggct	1560
atctcctatg	cccttgga	ggacagtgg	ctggggacag	gagaagtgt	caatcttcat	1620
gagacaaaag	ggcccgatca	aggcagccac	aaggccttga	cctgccgagt	cagcatgccc	1680
catctctctc	gacagctgtc	ccctaaaccc	aactcacgtt	tctgtatgtc	ttaggccagt	1740
atcccaaacc	tcttccacgt	cactgttctt	tccaccatt	ctccctttgc	atcttgagca	1800
gttatccaac	taggatctgc	caagtggata	ctggggtgcc	actcccctga	gaaaagactg	1860
agccaggaac	tacaagctcc	ccccacattc	ctccagcct	ggacctaat	cttgagaggg	1920
gctctctctt	cacggactgt	gtctggactt	tgagcaggct	tctgcccctt	gcgttggctc	1980
tttgctgcca	gccatcagg	gggggattag	agcctgggtg	aagtgcgcca	gactcttccg	2040
gtttccaaag	ttcgtgcctg	cgaacccaaa	cctgtgagtc	tcttctgcat	gcaggagttt	2100
ctcctgggca	gctggctcact	ccccagagaa	gctgggcctt	catggacaca	tggaaactaag	2160
cctoccaa	gggagttctg	gctgagccca	gggtggggag	atcctgggaa	gggagggcact	2220
ggaggaagac	ggcacctctt	cccccatggc	aggggtgtgag	ggaggcagg	ttggaatgg	2280
gcgagtatgg	caatctaagc	aggggtctgg	tctctttgac	tccaggctcg	ctttggccga	2340

```

ctgtctgctc acccagagac cttggactcc ggactatcca tggctccgaa tctaagtgtc 2400
gcccaactccc atgctcacac ccacagaagg tcttcccatc cccttttagat tctgtgctca 2460
ctccaccagt gaggaagatg cctctgtctt tcccacgact gccaggagat aggggaagccc 2520
agccaggact gacctcctt cctccagcct gccctgacct acctggcaaa gcagggcaca 2580
tggggaggaa gagactggaa cctttctttg acagccaggc ctagacagac aggcctgggg 2640
acactggccc atgaggggag gaaggcaggc gcacgaggtc caggagggcc cttttctgat 2700
catgcccctt ctctcccacc ccattctccc accaccacct ctgtggcctc catggtacct 2760
ccacagggtc ggcctcccct agagggtggg cctcaaccac ctctcccgc cagcaccgg 2820
ttagtggagc agggctgcca cgcaaccgcc aagccccct caaggtggga cagtacccc 2880
gacccatcca ctactcctg agaggctcgc gccagaatg ggaacctcag agaagagctc 2940
taaggagaag aaaccccata gcgtcagaga ggatatgtct ggcttccaag agaaaggagg 3000
ctccgttttg caaagtggag gagggacgag ggacaggggt ttcaccagcc agcaacctgg 3060
gccttgtagt gtctgtgttt ttaaaaccac taaagtgcaa gaattacatt gcactgtttc 3120
tccacttttt attttctctt aggtttttgt ttctatttca aacatacttt cttggttttc 3180
taatggagta tatagtttag tcatttcaca gactctggcc tctctcctg aaatcctttt 3240
ggatggggaa agggaagggt gggagggtcc gaggggaagg ggacccagc ttcctgtgtc 3300
cgcctcacc cactccacca gtcccgggtc gccagccgga gtctccttc taccgccact 3360
gtcacaccgt agccacatg gatagcacag ttgtcagaca agattccttc agattccgag 3420
ttgctaccgg ttgttttcgt tgttgttgtt gttgtttttc tttttctttt tttttttgaa 3480
gacagcaata accacagtac atattactgt agttctctat agttttacat acattcatac 3540
cataactctg ttctctcctc ttttttgttt tcaactttaa aaacaaaaat aaacgatgat 3600
aatctttact ggtgaaaagg atggaaaaat aaatcaacaa atgcaaccag tttgtgagaa 3660
aaaaaaaaaa aa 3672

```

<210> 30
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> L27560

<400> 30
 agcaacctgg gccttgtagt gtctgtgttt ttaaaaccac taaagtgcaa gaattacatt 60

<210> 31
 <211> 1416
 <212> DNA
 <213> Homo sapiens

<220>
 <221> Modified_base
 <222> 1 ... 1416
 <223> n = a,c,g, or t

<300>
 <308> M55914

<400> 31
 aggaattccg gaattccgga attccgatgg atggaacaga aaataaatct aagtttggtg 60
 cgaacgccat tctgggggtg tcccttgccg tctgcaaagc tggtgccgtt gagaaggggg 120
 tcccctgtac cgccacatcg cgtacttggc tggcaacttc gaagtcacatc tgccagtgccc 180
 ggcgttcaag tgtcatcatc aatggcgggt ctcatgctgg caacaagctg gccatgcaga 240
 gtctgtcctc ccagtcgggtg cagcaaactc aggggaagcca tgccgcattg gagcagagggt 300
 ttaccacaac ctgaagaatg tcatcaagga gaaatatggg aaagatgcca ccaatgtggg 360
 gatttgcgcg ggtttgctcc caacatcctg gagaataaag aaggcctgga gctgctgaag 420
 actgctattg gaaagcctgg cctacactgt aaagggtggc atggcatgga cgtagcggcc 480
 tccgagttct tcaggtcagg gaactatgac ctggacttca agtctcccga tgacccagc 540
 aggtacatct cgcctgacca gctggctgac ctgtacaagt ccttcacaa ggactaccca 600
 gtggtgtcta tcgaagatcc ctttgaccag gatgactggg gagcttcaga agttcacagc 660
 cagtgcagga atccaggtag tgggggggatg actcacagt accaacccaa agaggatcgc 720

```

caaggcgtga acgagaagtc ctgcaactgc ctccctgctca aagtcaacca gattgggtcc 780
gtgaccgagt ctcttcaggc gtgcaagctg gccaggcca atggttgggg cgtcatgggtg 840
tctcatcggt cgggggagac tgaagatacc ttcatcgctg acctgggtgt ggggctgtgc 900
actggggcag atcaagactg gtgccccttg ccgatcacgc gcttggccaa gtacaaccag 960
ctccctcagaa ttgaagagga gctgggcagc aaggctaagt ttgccggcag gaacttcaga 1020
aacccttgg ccaagtaagc tgtgggcagg caagccttcg gtcacctgtt ggctacagac 1080
ccctcccctg gtgtcagctc aggcagctcg aggccccga ccaacacttg caggggtccc 1140
tgctagttag cgccaccgc cgtggagttc gtaccgcttc cttagaactc tacagaagcc 1200
aagctccctg gaagccctgt tggcagctct agctttgcag ttgtgtaatt ggcccaagtc 1260
attgtttttc tcgccttact ttccaccaag tgtctagagt catgtgagcc tngtgtcatc 1320
tccggggtgg ccacaggta gatccccgtt ggttttgtgc tcaaaaataaa aagcctcagt 1380
gacccatgaa aaaaaaaaaa gaattccgga attccg 1416

```

<210> 32

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> M55914

<400> 32

```

gtaccgcttc cttagaactc tacagaagcc aagctccctg gaagccctgt tggcagctct 60

```

<210> 33

<211> 2517

<212> DNA

<213> Homo sapiens

<300>

<308> M96577

<400> 33

```

ggaattccgt ggccgggact ttgcaggcag cggcggccgg gggcggagcg ggatcgagcc 60
ctgcgcgagg cctgccgcca tgggcccgcg ccgcgcgcgc cgcctgtcac ccgggccgcg 120
cgggccgtga gcgtcatggc cttggccggg gcccctgagg ggggccatg ccgcgccggc 180
ctggaggccc tgctcggggc cggcgcgctg cggctgtctg actcctcgca gatcgatc 240
atctccgccc cgcaggagcg cagcgccccg ccggctccca ccggccccgc ggcgccccgc 300
gccggccctt gcgacctga cctgctgtct ttgccacac cgcaggcgcc ccggcccaca 360
cccagtgccc cgcggccccg gctcggccgc ccgccggtga agcggaggct ggacctggaa 420
actgaccatc agtacctggc cgagagcagt gggccagctc ggggcagagg ccgccatcca 480
ggaaaagggt tgaatatccc gggggagaag tcacgctatg agacctact gaatctgacc 540
accaagcgct tcctggagct gctgagccac tcggctgacg gtgtcgtcga cctgaactgg 600
gctgccgagg tgctgaaggt gcagaagcgg cgcactatg acatcaccaa cgtccttgag 660
ggcatccagc tcattgccaa gaagtccaag aaccacatcc agtggctggg cagccacacc 720
acagtgggcg tcggcggagc gcttgagggg ttgaccagg acctccgaca gctgcaggag 780
agcgagcagc agctggacca cctgatgaat atctgtacta cgcagctgcg cctgctctcc 840
gaggacatcg acagccagcg cctggcctac gtgacgtgtc aggaccttcg tagcattgca 900
gacctgcag agcagatggt tatggtgatc aaagcccctc ctgagaccca gctccaagcc 960
gtggactctt cggagaactt tcagatctcc cttagagca aacaaggccc gatcgatgtt 1020
ttcctgtgcc ctgaggagac cgtaggtggg atcagccctg ggaagacccc atcccaggag 1080
gtcacttctg aggaggagaa cagggccact gactctgcca ccatagtgtc accaccacca 1140
tcacttcccc cctcatccct caccacagat ccagccagt ctctactcag cctggagcaa 1200
gaaccgtgtg tgtcccggat gggcagcctg cgggctcccg tggacgagga ccgctgtcc 1260
ccgtgggtgg cggccgactc gctcctggag catgtgcggg aggacttctc cggcctcttc 1320
cctgaggagt tcatcagcct tccccaccc cacgaggccc tcgactacca cttcggcctc 1380
gaggagggcg agggcatcag agacctcttc gactgtgact ttggggacct cccccctg 1440
gatttctgac agggcttgga gggaccaggg tttccagagt agctcacctt gtctctgcag 1500
ccctggagcc ccctgtccct ggccgtcctc ccagcctgtt tggaaacatt taatttatac 1560
ccctctctct tgtctccaga agcttctagc tctggggtct ggctaccgct aggaggctga 1620

```

```

gcaagccagg aaggggaagga gtctgtgtgg tgtgtatgtg catgcagcct acaccacac 1680
gtgtgtaccg ggggtgaatg tgtgtgagca tgtgtgtgtg catgtaccgg ggaatgaagg 1740
tgaacataca cctctgtgtg tgcactgcag acacgcecca gtgtgtccac atgtgtgtgc 1800
atgagtcacat ctctgcgcgt ggggggggctc taactgcact ttcgccctt ttgctcgtgg 1860
ggccccacaa ggcccagggc agtgccctgt cccagaatct ggtgctctga ccaggccagg 1920
tggggaggct ttggctggct gggcgtgtag gacgggtgaga gcacttctgt cttaaagggt 1980
ttttctgatt gaagctttta tggagcgtta tttattttatc gaggcctctt tgggtgagcct 2040
ggggaatcag caaaaaggga ggaggggtgt ggggttgata ccccaactcc ctctaccctt 2100
gagcaagggc aggggtccct gagctgttct tctgccccat actgaaggaa ctgaggcctg 2160
ggtgatttat ttattgggaa agtgaggag ggagacagac tgactgacag ccatgggtgg 2220
tcagatggtg ggggtggccc tctccagggg gccagttcag ggcccagctg cccccagga 2280
tggatatgag atgggagagg tgagtggggg accttccactg atgtgggcag gagggggtgg 2340
gaaggcctcc cccagcccag accctgtggg cctcctgca gtgtctgaag cgctgcctc 2400
cccactgctc tgccccaccc tccaatctgc actttgattt gcttcctaac agctctgttc 2460
cctcctgctt tggttttaat aaatatattg atgacgttaa aaaaaggaat tcgatat 2517

```

<210> 34

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> M96577

<400> 34

```

gtaggacggt gagagcactt ctgtcttaaa ggttttttct gattgaagct ttaatggagc 60

```

<210> 35

<211> 4437

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000057

<400> 35

```

gcgcggcggc cgtgggttgcg gcgcgggaag tttggatcct ggttcogtcc gctaggagtc 60
tgcgtgcgag gattatggct gctgttcttc aaaataatct acaggagcaa ctagaacgtc 120
actcagccag aacacttaat aataaattaa gtctttcaca accaaaattt tcaggtttca 180
cttttaaaaa gaaaacatct tcagataaca atgtatctgt aactaatgtg tcagtagcaa 240
aaacacctgt attaagaaat aaagatgtta atgttacoga agacttttcc ttcagtgaac 300
ctctacccaa caccacaaat cagcaaaagg tcaaggactt ctttaaaaaat gctccagcag 360
gacaggaaac acagagaggt ggatcaaaat cattattgcc agattttctt cagactccga 420
aggaagtgtg atgcactacc caaaacacac caactgtaaa gaaatcccg gatactgctc 480
tcaagaaatt agaatttagt tcttcaccag attctttaag taccatcaat gattgggatg 540
atatggatga ctttgatact tctgagactt caaaatcatt tgttacacca ccccaaagtc 600
actttgtaag agtaagcact gctcagaaat caaaaaagg taagagaaac ttttttaag 660
cacagcttta tacaacaaac acagtaaaga ctgatttgcc tccacctcc tctgaaagcg 720
agcaaataga tttgactgag gaacagaagg atgactcaga atgggttaagc agcgatgtga 780
tttgcatcga tgatggcccc attgctgaag tgcataataa tgaagatgct caggaaagtg 840
actctctgaa aactcatttg gaagatgaaa gagataatag cgaaaagaag aagaatttgg 900
aagaagctga attacattca actgagaaag ttccatgtat tgaatttgat gatgatgatt 960
atgatacggg ttttgttcca ccttctccag aagaaattat ttctgcttct tcttctctct 1020
caaaatgcct tagtacgtta agggaccttg acacatctga cagaaaagag gatgttctta 1080
gcacatcaaa agatcttttg tcaaaacctg agaaaatgag tatgcaggag ctgaattccag 1140
aaaccagcac agactgtgac gctagacaga taagtttaca gcagcagctt attcatgtga 1200
tggagcacat ctgtaaatta attgatacta ttctgatga taaactgaaa cttttggatt 1260
gtgggaacga actgcttcag cagcgggaaca taagaaggaa acttctaacg gaagtagatt 1320
ttaataaaaag tgatgccagt cttcttggct cattgtggag atacaggcct gattcacttg 1380
atggccctat ggagggtgat tctgccccta cagggaattc tatgaaggag ttaaattttt 1440
cacaccttcc ctcaaattct gtttctcctg gggactgttt actgactacc accctaggaa 1500

```

```

agacaggatt ctctgccacc aggaagaatc tttttgaaag gcctttattc aatacccatt 1560
tacagaagtc cttttgtaagt agcaactggg ctgaaacacc aagactagga aaaaaaatg 1620
aaagctctta tttcccagga aatgtttctca caagcactgc tgtgaaagat cagaataaac 1680
atactgcttc aataaatgac ttagaaagag aaaccaacc ttcctatgat attgataatt 1740
ttgacataga tgactttgat gatgatgatg actgggaaga cataatgcat aatttagcag 1800
ccagcaaatc ttccacagct gcctatcaac ccatcaagga aggtcggcca attaaatcag 1860
tatcagaaag acttttctca gccaaagacag actgtcttcc agtgtcatct actgctcaaa 1920
atataaactt ctccagagtc attcagaatt atactgacaa gtcagcacia aatttagcat 1980
ccagaaatct gaaacatgag cgtttccaaa gtcttagttt tcctcatata aaggaaatga 2040
tgaagatttt tcataaaaaa tttggcctgc ataattttag aactaatcag cttagggcga 2100
tcaatgctgc actgcttggt gaagactggt ttatcctgat gccgactgga ggtggtaaga 2160
gtttgtgtta ccagctccct gcctgtggtt ctctgggggt cactgttggtc atttctccct 2220
tgagatcact tatcgtagat caagtccaaa agctgacttc cttggatatt ccagctacat 2280
atctgacagg tgataagact gactcagaag ctacaaatat ttacctccag ttatcaaaaa 2340
aagacccaat cataaaactt ctatatgtca ctccagaaaa gatctgtgca agtaacagac 2400
tcattttctac tctggagaat ctctatgaga ggaagctctt ggcacgtttt gttattgatg 2460
aagcacattg tgtcagtcag tggggacatg attttcgtca agattacaaa agaatagaata 2520
tgcttcgccca gaagtttcct tctgttccgg tgatggctct tacggccaca gctaatacca 2580
gggtacagaa ggacatcctg actcagctga agattctcag acctcaggtg ttagcatga 2640
gctttaacag acataatctg aaatactatg tattaccgaa aaagcctaaa aagggtggcat 2700
ttgattgcct agaatggatc agaaagcacc acccatatga ttcagggata atttactgcc 2760
tctccaggcg agaatgtgac acctggctg acacgttaca gagagatggg ctogctgctc 2820
ttgcttaccg tgctggcctc agtgattctg ccagagatga agtgcagcag aagtggatta 2880
atcaggatgg ctgtcaggtt atctgtgcta caattgcatt tgggaatgggg attgacaaac 2940
cggacgtgcg atttgtgatt catgcatctc tcctaaatc tgtggagggt tactaccaag 3000
aatctggcag agctggaaga gatggggaaa tatctcactg cctgcttttc tatacctatc 3060
atgatgtgac cagactgaaa agacttataa tgatggaaaa agatggaaac catcatacaa 3120
gagaaactca cttcaataat ttgtatagca tgggtacatta ctgtgaaaat ataacggaat 3180
gcaggagaat acagctttttg gcctactttg gtgaaaaatg atttaatcct gatttttgta 3240
agaaacaccc agatgtttct tgtgataatt tgtgataaac aaaggattat aaaacaagag 3300
atgtgactga cgtatgtaaa agtattgtaa gatttgttca agaacatagt tcatcacaag 3360
gaatgagaaa tataaaacat gtaggtcctt ctggaagatt tactatgaat atgctggtcg 3420
acattttctt ggggagtaag agtgcaaaaa tccagtcagg tatatttgga aaaggatctg 3480
cttattcacg acacaatgcc gaaagacttt ttaaaaagct gatacttgac aagattttgg 3540
atgaagactt atatatcaat gccaatgacc aggcgatcgc ttatgtgatg ctcgaaata 3600
aagcccaaac tgtactaaat ggcaatttaa aggtagactt tatggaaaca gaaaattcca 3660
gcagtgtgaa aaaacaaaaa gcgttagtag caaaagtgtc tcagagggaa gagatgggta 3720
aaaaatgtct tggagaactt acagaagtct gcaaactctc ggggaaagtt tttgggtgtc 3780
attacttcaa tatttttaat accgtcactc tcaagaagct tgcagaatct ttatcttctg 3840
atcctgaggt tttgcttcaa attgatgggt ttactgaaga caaactggaa aaatatgggtg 3900
cggaagtgat ttcagtatta cagaaatact ctgaaatggac atcgccagct gaagacagtt 3960
ccccagggat aagcctgtcc agcagcagag gcccggaag aagtgccgct gaggagcttg 4020
acgaggaaat acccgatatc tcccactact ttgcaagtaa aaccagaaat gaaaggaaga 4080
ggaaaaagat gccagcctcc caaaggctca agaggagaaa aactgcttcc agtgggtcca 4140
aggcaaaggg ggggtctgcc acatgtagaa agatatcttc caaaacgaaa tcctccagca 4200
tcattggatc cagttcagcc tcacatactt ctcaagcgac atcaggagcc aatagcaaat 4260
tggggattat ggctccaccg aagcctataa atagaccgtt tcttaagcct tcatatgcat 4320
tctcataaca accgaatctc aatgtacata gaccctcttt cttgtttgtc agcatctgac 4380
catctgtgac tataaagctg ttattcttgt tataccaaaa aaaaaaaaaa aaaaaaa 4437

```

<210> 36

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000057

<400> 36

taagccttca tatgcattct cataacaacc gaatctcaat gtacatagac cctctttctt 60

<210> 37
 <211> 2016
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000060

<400> 37
 gccagctgga gcgttttcgg ggctgtaaag ggagaatggc gcatgcgcac attcagggcg 60
 gaaggcgcg taagagcaga tttgtgggtc gcattatgtc tggagccaga agtaagcttg 120
 ctcttttccct ctgcggctgt tacgtgggtg ccctggggagc ccacaccggg gaggagagcg 180
 tggctgacca tcacgaggct gaatattatg tggctgccgt gtatgagcat ccatccatcc 240
 tgagtctgaa ccctctggct ctcatcagcc gccaaaggcg cttggagctc atgaaccaga 300

 accttgacat ctatgaacag caagtgatga ctgcagccca aaaggatgta cagattatag 360
 tgtttccaga agatggcatt catggattca actttacaag aacatccatt tatccatttt 420
 tggacttcac gccgtctccc caggtgggtc ggtggaaccc atgcctggag cctcacccgt 480
 tcaatgacac agagggtgctc cagcgctga gttgtatggc catcagggga gatattgtct 540
 tgggtggccaa tcttgggaca aaggagcctt gtcatagcag tgaccacaag tgcccaaaag 600
 atgggagata ccagttcaac acaaagtgcg tgttcagcaa taatggaacc cttgttgacc 660
 gctaccgtaa acacaacctc tactttgagg cagcattcga tgttcctctt aaagtggatc 720
 tcatcacctt tgataccccc tttgctggca ggtttggcat cttcacatgc tttgatata 780
 tgttctttga ccctgccatc agagtcctca gagactacaa ggtgaagcat gttgtgtacc 840
 caactgcctg gatgaaccag ctcccactct tggcagcaat tgagattcag aaagcttttg 900
 ctgttgccct tggcatcaac gttctggcag ctaatgtcca ccaccagtt ctggggatga 960
 caggaagtgg catacacacc cctctggagt ccttttggtg ccatgacatg gaaaatccca 1020
 aaagtcacct tataattgcc caggtggcca aaaatccagt gggctctcatt ggtgcagaga 1080
 atgcaacagg tgaacaggac ccacccata gtaagttttt aaaaattttg tcaggcgatc 1140
 cgtactgtga gaaggatgct caggaagtcc actgtgatga ggccaccaag tggaaagtga 1200
 atgtctctcc cacatttcac tctgagatga tgtatgacaa tttcacccctg gtccctgtct 1260
 ggggaaagga aggttatctc cagctctgtt ccaatggcct ctgctgttat ttactttacg 1320
 agaggccac cttatccaaa gagctgtatg ccctgggggt ctttgatggg cttcacacag 1380
 tacatggcac ttactacatc caagtgtgtg ccctgggtcag gtgtgggggt cttggcttcg 1440
 acacctgagg acaggaaatc acagaggcca cggggatatt tgagtttcac ctgtggggca 1500
 acttcagtag ttcttatatc tttcctttgt ttctgacctc agggatgacc ctagaagtcc 1560
 ctgaccagct tggctgggag aatgaccact atttcctgag gaaaagttag ctgtcctctg 1620
 ggctgggtgac ggcggtctc tatgggcgt tgtatgagag ggactaggaa aagtgtgtgg 1680
 tctgtggggc ggactctggc catcatgttg acagccttgc acttcacag gctacaagcc 1740
 ctgggaccat ctttctgcct taagggcagg agccacttc tgtggacca gattccaccc 1800
 tgggaactgt ggaaaaagta ggagaggcag attccctcag tgtcttctc ttaaacctca 1860
 atcatcgaga cattaggggg tattttctgt tcacatttat ctttttcaag ccacatcttc 1920
 ctctaacaaa tctctcagta tgcgattggg ctcaagctaa aacaaaaata aatgtcagtt 1980
 tatattttac acatccaaaa aaaaaaaaaa aaaaaa 2016

<210> 38
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000060

<400> 38
 tcctctaaca aatctctcag tatgcgattg gtctcaagct aaaacaaaaa taaatgtcag 60

<210> 39
 <211> 811
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_000269

<400> 39

```

gcagaagcgt tccgtgcgtg caagtgcgtc gaaccacgtg ggtcccgggc gcgtttcggg 60
tgctggcggc tgcagccgga gttcaaacct aagcagctgg aaggaaccat ggccaactgt 120
gagcgtacct tcattgcgat caaaccagat ggggtccagc ggggtcttgt gggagagatt 180
atcaagcgtt ttgagcagaa aggattccgc cttgttggtc tgaaattcat gcaagcttcc 240
gaagatcttc tcaaggaaca ctacgttgac ctgaaggacc gtccattctt tgccggcctg 300
gtgaaataca tgcactcagg gccggtagtt gccatggtct gggaggggct gaatgtggtg 360
aagacggggc gagtcatgct cggggagacc aaccctgcag actccaagcc tggggaccatc 420
cgtggagact tctgcataca agttggcagg aacattatac atggcagtga ttctgtggag 480
agtgcagaga aggagatcgg cttgtggttt caccctgagg aactggtaga ttacacgagc 540
tgtgtctcaga actggatcta tgaatgacag gagggcagac cacattgctt ttcacatcca 600
tttccctcc ttcccatggg cagaggacca ggctgtagga aatctagtta ttacaggaa 660
cttcatcata atttggaggg aagctcttgg agctgtgagt tctccctgta cagtgttacc 720
atccccgacc atctgattaa aatgcttcct ccagcatag gattcattga gttggttact 780
tcatattggt gcattgcttt ttttccctt t 811

```

<210> 40

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000269

<400> 40

```

gtctgaaatt catgcaagct tccgaagatc ttctcaagga aactacgtt gacctgaagg 60

```

<210> 41

<211> 2338

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000291

<400> 41

```

agcgcacgtc ggcagtcggc tccctcggtg accgaatcac cgacctctct cccagctgt 60
atttccaaaa tgtcgttttc taacaagctg acgctggaca agctggacgt taaaggggaag 120
cgggtcggtta tgagagtcga cttcaatggt cctatgaaga acaaccagat acaaaacaac 180
cagaggatta aggctgctgt cccaagcatc aaattctgct tggacaatgg agccaagtgc 240
gtagtcctta tgagccacct aggcgggcct gatgggtgtg ccatgcctga caagtactcc 300
ttagagccag ttgctgtaga actcaaatct ctgctgggca aggatgttct gttcttgaag 360
gactgtgtag gcccagaagt ggagaaagcc tgtgccaacc cagctgctgg gtctgtcatc 420
ctgctggaga acctccgctt tcatgtggag gaagaaggga agggaaaaga tgcttctggg 480
aacaagggtta aagccgagcc agccaaaata gaagctttcc gagcttcaact ttccaagcta 540
gggatgtct atgtcaatga tgcttttggc actgctcaca gagcccacag ctccatggta 600
ggagtcaatc tgccacagaa ggctggtggg tttttgatga agaaggagct gaactacttt 660
gcaaaggcct tggagagccc agagcgaccc ttcctggcca tcctgggcgg agctaaagtt 720
gcagacaaga tccagctcat caataatatg ctggacaaag tcaatgagat gattattggg 780
ggtggaatgg cttttacctt ccttaagggt ctcaacaaca tggagattgg cacttctctg 840
tttgatgaag agggagccaa gattgtcaaa gacctaattg ccaaagctga gaagaatggg 900
gtgaagatta ccttgctgtg tgactttgtc actgctgaca agtttgatga gaatgccaaag 960
actggccaag ccactgtggc ttctggcata cctgctggct ggatgggctt ggactgtggg 1020
cctgaaagca gcaagaagta tgctgaggct gtcactcggg ctaagcagat tgtgtggaat 1080
ggctctgtgg gggatatttg atgggaagct tttgcccggg gaaccaaagc tctcatggat 1140
gaggtggtga aagccacttc taggggctgc atcaccatca taggtggtgg agacactgcc 1200
acttgctgtg ccaaatggaa cacggaggat aaagtcagcc atgtgagcac tgggggtggg 1260
gccagtttgg agctcctgga aggtaaagtc cttcctgggg tggatgctct cagcaatatt 1320

```

```

tagtactttc ctgcctttta gttcctgtgc acagccccta agtcaactta gcattttctg 1380
catctccact tggcattagc taaaaccttc catgtcaaga ttcagctagt ggccaagaga 1440
tgcagtgcc a ggaaccctta aacagttgca cagcatctca gctcatcttc actgcaccct 1500
ggatttgcat acattcttca agatcccatt tgaatttttt agtgactaaa ccattgtgca 1560
ttctagagtg catatattta ttttttgctt gttaaaaaga aagtgagcag tgtagctta 1620
gttctctttt gatgtagggt attatgatta gctttgtcac tgtttacta ctcagcatgg 1680
aaacaagatg aaattccatt tgtaggtagt gagacaaaat tgatgatcca ttaagtaaac 1740
aataaaagtg tccattgaaa ccgtgatttt tttttttttt ctgtcatact ttgttaggaa 1800
gggtgagaat agaattcttg ggaacggatc agatgtctat attgctgaat gcaagaagtg 1860
gggcagcagc agtggagaga tgggacaatt agataaatgt ccattcttta tcaagggcct 1920

```

```

actttatggc agacattgtg ctagtgtctt tattctaact tttattttta tcagttacac 1980
atgatcataa tttaaaaagt caaggcttat aacaaaaaag cccagccca ttcctcccat 2040
tcaagattcc cactccccag aggtgaccac tttcaactct tgagtttttc aggtatatac 2100
ctccatgttt ctaagtaata tgcttatatt gttaacttcc ttttttttta ttttttaaag 2160
aaatctattt cataccatgg aggaaggctc tgttccacat atatttccac ttcttcattc 2220
tctcgtata gttttgtcac aattatagat tagatcaaaa gtctacataa ctaatacagc 2280
tgagctatgt agtatgtat gattaaattt acttatgtaa aaaaaaaaaa aaaaaaaaaa 2338

```

<210> 42
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000291

<400> 42
 acttagcatt ttctgcatct ccacttggca ttagctaaaa ccttccatgt caagattcag 60

<210> 43
 <211> 787
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000363

```

<400> 43
ctgaaggcca cccgggcggc cccctcactg accctccaaa cggccctgtc ctgcgcttgc 60
ctcctgccat tcccggcctg agtctcagca tggcggatgg gagcagcgat gcggctaggg 120
aacctcgccc tgcaccagcc ccaatcagac gccgtccctc caactaccgc gcttatgcc 180
cggagccgca cgccaagaaa aaatctaaga tctccgctc gagaaaattg cagctgaaga 240
ctctgctgct gcagattgca aagcaagagc tggagcgaga ggcggaggag cggcgcgagg 300
agaaggggcy cgtcttgagc acccgctgcc agccgctgga gttgaccggg ctgggcttcg 360
cggagctgca ggacttgtgc cgacagctcc acgcccgtgt ggacaagggt gatgaagaga 420
gatacgacat agaggcaaaa gtcaccaaga acatcacgga gattgcagat ctgactcaga 480
agatctttga ccttcgaggc aagttaagc ggcccaccct gcggagagtg aggatctctg 540
cagatgccat gatgcaggcg ctgctggggg cccgggctaa ggagtccttg gacctgcggg 600
cccacctcaa gcaggagaag aaggaggaca ccgagaagga aaaccgggag gtgggagact 660
ggcggaagaa catcgatgca ctgagtggaa tggaggccg caagaaaaag tttgagagct 720
gagccttctt gctactgcc cctgccctga ggaggccac tgagggaataa agcttctctc 780
tgagctg 787

```

<210> 44
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000363

<400> 44
tgtgggacaag gtgggatgaag agagatacga catagaggca aaagtcacca agaacatcac 60

<210> 45
<211> 1263
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000365

<400> 45
ggcacgagac cttcagcgcc tgggctccag cgccatggcg ccctccagga agttcttctgt 60
tgggggaaac tgggaagatga acgggcgga gacagagtctg ggggagctca tcggcactct 120
gaacgcggcc aagggtgccgg ccgacaccga ggtgggttctg gctcccccta ctgcctatat 180
cgacttcgcc cggcagaagc tagatcccaa gattgctgtg gctgcgcaga actgctacaa 240
agtgaactaat ggggctttta ctggggagat cagccctggc atgatcaaag actgcggagc 300
cacgtgggtg gtccctggggc actcagagag aaggcatgtc tttggggagt cagatgagct 360
gattgggacg aaagtggccc atgctctggc agagggactc ggagtaatcg cctgcattgg 420
ggagaagcta gatgaaaggg aagctggcat cactgagaag gttgttttctg agcagacaaa 480
ggtcatcgca gataacgtga aggactggag caaggctcgtc ctggcctatg agcctgtgtg 540
ggccattggt actggcaaga ctgcaacacc ccaacaggcc caggaagtac acgagaagct 600
ccgaggatgg ctgaagtcca acgtctctga tgcggtggct cagagcacc gtatcattta 660
tggaggctct gtgactgggg caacctgcaa ggagctggcc agccagcctg atgtggatgg 720
cttccttctg ggtgggtgctt ccctcaagcc cgaattcgtg gacatcatca atgccaaca 780
atgagcccca tccatcttcc ctacccttcc tgccaagcca gggactaagc agcccagaag 840
cccagtaact gccctttccc tgcataatgct tctgatgggtg tcatctgtctc ctccctgtgg 900
cctcatccaa actgtatctt cctttactgt ttatatcttc accctgtaat ggttgggacc 960
aggccaatcc ctctccact tactataatg gttggaacta aacgtcacca aggtggcttc 1020
tccttggctg agagatggaa ggcgtgggtg gatttgcctc tgggttccct agggccctagt 1080
gagggcagaa gagaaacat cctctccctt cttacaccgt gaggccaaaga tcccctcaga 1140
aggcaggagt gctgccctct cccatgggtgc ccgtgcctct gtgctgtgta tgtgaaccac 1200
ccatgtgagg gaataaacct ggcactagga aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1260
aaa 1263

<210> 46
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000365

<400> 46
tatcttcacc ctgtaatggt tgggaccagg ccaatccctt ctccacttac tataatgggtt 60

<210> 47
<211> 1616
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000582

<400> 47
ctccctgtgt tgggtggagga tgtctgcagc agcattttaa ttctgggagg gcttgggtgt 60
cagcagcagc aggaggaggc agagcacagc atcgctggga ccagactcgt ctcaggccag 120
ttgcagcctt ctgagccaaa cgccgaccaa ggaaaactca ctaccatgag aattgcagtg 180
atttgccttt gcctcctagg catcacctgt gccataccag ttaaacaggc tgattctgga 240
agttctgagg aaaagcagct ttacaacaaa taccagatg ctgtggccac atggctaaac 300

```

cctgacccat ctcagaagca gaatctccta gccccacaga cccttccaag taagtccaac 360
gaaagccatg accacatgga tgatatggat gatgaagatg atgatgacca tgtggacagc 420
caggactcca ttgactcgaa cgactctgat gatgtagatg acactgatga ttctcaccag 480
tctgatgagt ctcaccattc tgatgaatct gatgaactgg tcaactgattt tcccacggag 540
ctgccagcaa ccgaagtttt cactccagtt gtccccacag tagacacata tgatggccga 600
gggtgatagt tggtttatgg actgagggtca aaatctaaga agtttcgcag acctgacatc 660
cagtaccctg atgctacaga cgaggacatc acctcacaca tggaaagcga ggagtgaat 720
ggtgcataca aggccatccc cgttgccag gacctgaacg cgccttctga ttgggacagc 780
cgtgggaagg acagttatga aacgagtcag ctggatgacc agagtgcctga aaccacagc 840
cacaagcagt ccagattata taagcggaag gccaatgatg agagcaatga gcattccgat 900
gtgattgata gtcaggaact ttccaaagtc agccgtgaat tccacagcca tgaatttcac 960
agccatgaag atatgctggg tgtagacccc aaaagtaagg aagaagataa acacctgaaa 1020
tttcgtatatt ctcattgaatt agatagtgc tcttctgagg tcaattaaaa ggagaaaaaa 1080
tacaatttct cactttgcat ttagtcaaaa gaaaaaatgc tttatagcaa aatgaaagag 1140
aacatgaaat gcttctttct cagtttattg gttgaatgtg tatctatttg agtctggaaa 1200
taactaatgt gtttgataat tagtttagtt tgtggcttca tggaaactcc ctgtaaacta 1260
aaagcttcag gggtatgtct atgttcattc tatagaagaa atgcaaacta tcaactgtatt 1320
ttaatatattg ttattctctc atgaatagaa atttatgtag aagcaaacia aatactttta 1380
cccacttaaa aagagaatat aacattttat gtcactataa tcttttgttt ttttaagttag 1440
tgtataatttt gttgtgatta tctttttgtg gtgtgaataa atcttttatc ttgaatgtaa 1500
taagaatttg gtggtgtcaa ttgcttattt gttttccac ggttgtccag caattaataa 1560
aacataacct tttttactgc ctaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1616

```

<210> 48

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000582

<400> 48

```

gggtggtgtca attgcttatt tgttttccca cggttgtcca gcaattaata aaacataacc 60

```

<210> 49

<211> 1666

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000584

<400> 49

```

ctccataagg cacaaacttt cagagacagc agagcacaca agcttctagg acaagagcca 60
ggaagaaacc accggaagga accatctcac tgtgtgtaaa catgaacttc aagctggccg 120
tggctctctt ggcagccttc ctgatttctg cagctctgtg tgaagggtgca gttttgccaa 180
ggagtgtctaa agaacttaga tgtcagtgc taaagacata ctccaaacct ttccacccca 240
aatttatcaa agaactgaga gtgattgaga gtggaccaca ctgcgccaac acagaaatta 300
ttgtaaagct ttctgatgga agagagctct gtctggacc caaggaaaac tgggtgcaga 360
gggttgtgga gaagtgtttg aagagggtct agaattcata aaaaaattca ttctctgtgg 420
tatccaagaa tcagtgaaga tgccagtga acttcaagca aatctacttc aacacttcat 480
gtattgtgtg ggtctgttgt aggggttgcca gatgcaatac aagattcctg gttaaatttg 540
aatttcagta aacaatgaat agtttttcat tgtaccatga aatatccaga acatacttat 600
atgtaaagta ttattttatt gaatctacaa aaaacaacia ataattttta aatataagga 660
ttttcctaga tattgcacgg gagaatatac aaatagcaaa attgaggcca agggccaaga 720
gaatatccga actttaattt caggaattga atgggtttgc tagaatgtga tatttgaagc 780
atcacataaa aatgatggga tgccataaag tcaaatttag ctggaaatcc 840
tggatttttt tctgttaaat ctggcaaccc tagtctgcta gccaggatcc acaagtcctt 900
gttccactgt gccttggttt ctcttttatt tctaagtgga aaaagtatta gccaccatct 960
tacctcacag tgatgttgtg aggacatgtg gaagcacttt aagttttttc atcataacat 1020
aaattatttt caagtgtaac ttattaacct atttattatt tatgtattta ttttaagcatc 1080

```

```

aaatatattgt gcaagaattt ggaaaaatag aagatgaatc attgattgaa tagttataaa 1140
gatgtttatag taaatattatt ttatttttaga tattaaatga tgttttatta gataaatttc 1200
aatcaggggtt tttagattaa acaaaacaaac aattgggtac ccagttaaat tttcatttca 1260
gataaacaac aaataatttt ttagtataag tacattattg tttatctgaa attttaattg 1320
aactaacaat cctagtttga tactcccagt cttgtcattg ccagctgtgt tggtagtgct 1380
gtgttgaatt acggaataat gagttagaac tattaaaaca gccaaaactc cacagtcaat 1440
attagtaatt tcttgctggg tgaaacttgt ttattatgta caaatagatt cttataatat 1500
tatttaaattg actgcatttt taaatacaag gctttatatt ttttaacttta agatgttttt 1560
atgtgctctc caaatttttt ttactgtttc tgattgtatg gaaatataaa agtaaatatg 1620
aaacattttaa aatataattt gttgtcaaag taaaaaaaaa aaaaaa 1666

```

```

<210> 50
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_000584

```

```

<400> 50
tggtagtgt gtgttgaatt acggaataat gagttagaac tattaaaaca gccaaaactc 60

```

```

<210> 51
<211> 1722
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_000599

```

```

<400> 51
ggggaaaaaga gctaggaaaag agctgcaaag cagtgtgggc tttttccctt tttttgctcc 60
ttttcattac cctcctcccg ttttcacctt tctccggact tcgcgtagaa cctgcgaatt 120
tcgaagagga ggtggcaaaag tgggagaaaa gaggtgttag ggtttggggg tttttgtttt 180
ttgtttttgt tttttaattt cttgatttca acattttctc ccacctctc ggctgcagcc 240
aacgcctctt acctgtttct cggcgccgcg caccgctggc agctgagggg tagaaagcgg 300
ggtgtatttt agattttaag caaaaatttt aaagataaat ccatttttct cttccacccc 360
caacgccatc tccactgcat ccgatctcat tatttcgggtg gttgcttggg ggtgaacaat 420
tttgtggctt tttttccctt ataattctga cccgctcagg cttgaggggt tctccggcct 480
cgcctcactg cgtgcacctg gcgctgcctt gcttccccca acctgttgca aggccttaat 540
tcttgcaact gggacctgct cgcaggcacc ccagccctcc acctctctct acatttttgc 600
aagtgtctgg gggagggcac ctgctctacc tgccagaaat tttaaaacaa aaacaaaaac 660
aaaaaaatct ccggggggccc tcttggtccc tttatccctg cactctcgtc ctctgcccc 720
accccgaggt aaagggggcg actaagagaa gatggtgttg ctaccgcgg tcctcctgct 780
gctggccgcc tatgccccgg cggcccagag cctgggctcc ttcgtgcact gcgagccctg 840
cgacgagaaa gccctctcca tgtgcccccc cagccccctg ggctgcgagc tggtaacgga 900
gccgggctgc ggctgctgca tgacctgcgc cctggccgag gggcagtcgt gcggcgtcta 960
caccgagcgc tgcgcccagg ggctgcgctg cctcccccg caggacgagg agaagccgct 1020
gcacgccttg ctgcacggcc gcggggtttg cctcaacgaa aagagctacc gcgagcaagt 1080
caagatcgag agagactccc gtgagcacga ggagcccacc acctctgaga tggccgagga 1140
gacctactcc cccaagatct tccggcccaa acacaccgc atctccgagc tgaaggctga 1200
agcagtgaag aaggaccgca gaaagaagct gaccagtc aaagttgtcg ggggagccga 1260
gaacactgcc cacccccga tcatctctgc acctgagatg agacaggagt ctgagcaggg 1320
ccctgcccgc agacacatgg aggcctccct caggagctc aaagccagcc cagcatggt 1380
gccccgtgct gtgtacctgc ccaatttgtg ccgcaaagga ttctacaaga gaaagcagtg 1440
caaaccttcc cgtggccgca agcgtggcat ctgctgggtg gtggacaagt acgggatgaa 1500
gctgccaggc atggagtagc ttgacgggga ctttcagtgc cacacctctg acagcagcaa 1560
cgttgagtga tgcgtcccc ccacacctt cctcacccc ctcccacccc cagccccgac 1620
tccagccagc gcctccctcc accccaggac gccactcatt tcatctcatt taaggggaaaa 1680
atatatatct atctatttga ggaaaaaaaa aaaaaaaaaa aa 1722

```

<210> 52
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000599

<400> 52
 ccaggacgcc actcatttca tctcatttaa gggaaaaata tatatctatc tatttgagga 60

<210> 53
 <211> 704
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000735

<400> 53
 gcagttactg agaactcata agacgaagct aaaatccctc ttccgatcca cagtcaaccg 60
 ccctgaacac atccctgcaaa aagcccagag aaaggagcgc catggattac tacagaaaat 120
 atgcagctat ctttctgggc acattgtcgg tgtttctgca tgttctccat tccgctcctg 180
 atgtgcagga ttgccagaa tgcacgctac aggaaaacc attcttctcc cagccgggtg 240
 cccaatact tcagtgcctg ggctgctgct tctctagagc atatccact ccactaaggt 300
 ccaagaagac gatgttggtc caaaagaacg tcacctcaga gtccacttgc tgtgtagcta 360
 aatcatataa cagggtcaca gtaatggggg gtttcaaagt ggagaaccac acggcgtgcc 420
 actgcagtac ttgttattat cacaaatctt aaatgtttta ccaagtgtg tcttgatgac 480
 tgcctgattt ctggaatgga aaattaagtt gtttagtggt tatggctttg tgagataaaa 540
 ctctcctttt ccttaccata ccactttgac acgcttcaag gatatactgc agctttactg 600
 ccttcctcct taccctacag tacaatcagc agtctagtgc ttttcatttg gaatgaatac 660
 agcatlaagc ttgttccact gcaaataaag ccttttaaat catc 704

<210> 54
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000735

<400> 54
 tgagataaaa ctctcctttt ccttaccata ccactttgac acgcttcaag gatatactgc 60

<210> 55
 <211> 1342
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000799

<400> 55
 cccggagccg gaccggggcc accgcgcccg ctctgctccg acaccgcgcc ccctggacag 60
 ccgccctctc ctccaggccc gtggggctgg ccctgcaccg ccgagcttcc cgggatgagg 120
 gcccccggtg tggtcacccg gcgcgccccca ggctcgctgag ggaccccggc caggcgcgga 180
 gatgggggtg cacgaatgtc ctgcctggct gtggcttctc ctgtccctgc tgtcgctccc 240
 tctgggcctc ccagtcctgg gcgccccacc acgcctcctc tgtgacagcc gaggcctgga 300
 gaggtacctc ttggaggcca aggaggccga gaatatcagc acgggctgtg ctgaacactg 360
 cagcttgaat gagaatatca ctgtcccaga caccaaagtt aatttctatg cctggaagag 420

```

gatggagggtc gggcagcagg ccgtagaagt ctggcagggc ctggccctgc tgtcgggaagc 480
tgtcttcgcg ggccaggccc tgttgggtcaa ctcttcccag ccgtgggagc ccctgcagct 540
gcatgtggat aaagccgtca gtggcccttcg cagcctcacc actctgcttc gggctctgcg 600
agcccagaag gaagccatct cccctccaga tggggcctca gctgctccac tccgaacaat 660
cactgctgac actttccgca aactcttccg agtctactcc aatttccctc ggggaaagct 720
gaagctgtac acaggggagg cctgcaggac aggggacaga tgaccagggtg tgtccacctg 780
ggcataatcca ccacctccct caccaacatt gcttgtgcca caccctcccc cgccactcct 840
gaaccccgtc gaggggctct cagctcagcg ccagcctgtc ccattggacac tccagtgcga 900
gcaatgacat ctcaaggggc agagggaactg tccagagagc aactctgaga tctaaggatg 960
tcacagggcc aacttgaggg cccagagcag gaagcattca gagagcagct ttaaactcag 1020
ggacagagcc atgctgggaa gacgcctgag ctactcggc accctgcaaa atttgatgcc 1080
aggacacgct ttggaggcga tttacctgtt ttgcaccta ccatcaggga caggatgacc 1140
tggaagaactt aggtggcaag ctgtgacttc tccagggtct acgggcatgg gcactccctt 1200
ggtggcaaga gcccccttga caccgggggtg gtgggaacca tgaagacagg atgggggctg 1260
gcctctggct ctcatgggg ccaagttttg tgtattcttc aacctcattg acaagaactg 1320
aaaccaccaa aaaaaaaaaa aa 1342

```

<210> 56
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000799

```

<400> 56
tcattgggggtc caagttttgt gtattcttca acctcattga caagaactga aaccacaaaa 60

```

<210> 57
 <211> 2722
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000917

```

<400> 57
gagcgggctg agggtaggaa gtagccgctc cgagtggagg cgactggggg ctgaagagcg 60
cgccgccctc tcgtcccact ttccagggtgt gtgatcctgt aaaattaaat cttccaagat 120
gatctgggat atattaatta taggaattct gcttcccag tctttggctc atccaggctt 180
ttttacttca attggtcaga tgactgattt gatccatact gagaaagatc tgggtgacttc 240
tctgaaagat tatattaagg cagaagagga caagttagaa caaataaaaa aatgggcaga 300
gaagttagat cggctaacta gtacagcgac aaaagatcca gaaggatttg ttgggcatcc 360
agtaaatgca ttcaaattaa tgaaacgtct gaatactgag tggagtgagt tggagaatct 420
ggtccttaag gatatgtcag atggccttat ctctaacctc accattcaga gaccagtact 480
ttctaattgat gaagatcagg ttggggcagc caaagctctg ttacgtctcc aggataccta 540
caatttggat acagatacca tctcaaagg taatcttcca ggagtgaac acaaatcttt 600
tctaacggct gaggactgct ttgagttggg caaagtggcc tatacagaag cagattatta 660
ccatacggaa ctgtggatgg aacaagccct aaggcaactg gatgaaggcg agatttctac 720
catagataaa gtctctgttc tagattatct gagctatgag gtatatcagc agggagacct 780
ggataaggca cttttgctca caaagaagct tcttgaacta gatcctgaac atcagagagc 840
taatggtaac ttaaaatatt ttgagtatat aatggctaaa gaaaaagatg tcaataagtc 900
tgcttcagat gaccaatctg atcagaaaac tacaccaaag aaaaaagggg ttgctgtgga 960
ttacctgcca gagagacaga agtcaaaagt gctgtccgt ggggagggta tcaaatcttt 1020
ccctcggaga cagaaaaaac tcttttgccg ctaccatgat ggaaaccgta atcctaaatt 1080
tattctggct ccagctaaac aggaggatga atgggacaag cctcgtatta ttctgttcca 1140
tgatattatt tctgatgcag aaattgaaat cgtcaaagac ctacgcaaac caaggctgag 1200
ccgagctaca gtacatgacc ctgagactgg aaaattgacc acagcacagt acagagtatc 1260
taagagtgcc tggctctctg gctatgaaaa tctgtgggtg tctcgaatta atatgagaat 1320
acaagatcta acaggactag atgtttccac agcagaggaa ttacaggtag caaattatgg 1380
agttggagga cagtatgaac cccatthttga ctttgcacgg aaagatgagc cagatgcttt 1440

```

```

caaagagctg gggacaggaa atagaattgc tacatggctg ttttatatga gtgatgtgtc 1500
tgcaggagga gccactgttt ttcctgaagt tggagctagt gtttggccca aaaaaggaac 1560
tgctgttttc tgggtataatc tggttgccag tggagaagga gattatagta caggcatgc 1620
agcctgtcca gtgctagttg gcaacaaatg ggtatccaat aaatggctcc atgaacgtgg 1680
acaagaatct cgaagacctt gtacgttgtc agaattggaa tgacaaacag gcttcccttt 1740
ttctcctatt gttgtactct tatgtgtctg atatacacat ttccatagtc ttaactttca 1800
ggagtttaca attgactaac actccatgat tgattcagtc atgaacctca tcccatgttt 1860
catctgtgga caattgctta ctttgtgggt tcttttaaaa gtaacacgaa atcatcatat 1920
tgcataaaac cttaaagtto tggttggtatc acagaagaca aggcagagtt taaagtgagg 1980
aattttataat ttaaagaact ttttgggttg ataaaaacat aatttgagca tccagtttta 2040
gtatttcact acatctcagt tgggtgggtg taagctagaa tgggctgtgt gataggaaac 2100
aaatgcctta cagatgtgcc taggtgttct gtttacctag tgtcttactc tgttttctgg 2160

```

```

atctgaagac tagtaataaa ctaggacact aactgggttc catgtgattg ccctttcata 2220
tgatcttcta agttgatttt tttcctccca agtctttttt aaagaaagta tactgtattt 2280
taccaacccc ctctcttttc ttttagctcc tctgtggtga attaaacgta cttgagttta 2340
aatatttcga tttttttttt ttttttaatt gaaagtcctg cataacaaca ctgggccttc 2400
ttaactaaaa tgctcaccac ttagcctggt tttttatccc ttttttaaaa tgacagatga 2460
ttttgttcag gaattttgct gtttttctta gtgctaatac cttgcctctt attcctgcta 2520
cagcagggtg gtaatatggg cattctgatt aaatactgtg ccttaggaga ctggaagttt 2580
aaaaatgtac aagtcctttc agtgatgagg gaattgattt tttttaaaag tctttttctt 2640
agaaagccaa aatgtttgtt tttttaagat tctgaaatgt gttgtgacaa caatgacctt 2700
tttatgatct taaatctttt tt 2722

```

<210> 58
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000917

```

<400> 58
tcttactctg ttttctggat ctgaagacta gtaataaaact aggacactaa ctgggttcca 60

```

<210> 59
 <211> 3236
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001109

```

<400> 59
gacccggcca tgcgcggcct cgggctctgg ctgctgggcg cgatgatgct gcctgcgatt 60
gccccagcc ggccctgggc cctcatggag cagtatgagg tcgtgttgcc gcggcgctctg 120
ccaggccccc gagtccgccg agctctgccc tcccacttgg gcctgcaccc agagaggggtg 180
agctacgtcc ttggggccac agggcacaac ttcacctcc acctgcggaa gaacagggac 240
ctgctggggt ccggctacac agagacctat acggctgcca atggctccga ggtgacggag 300
cagcctcgcg ggcaggacca ctgcttatac cagggccacg tagaggggta cccggactca 360
gccgccagcc tcagcacctg tgccggcctc aggggtttct tccaggtggg gtcagacctg 420
cacctgatcg agcccctgga tgaaggtggc gagggcggac ggcacgccgt gtaccaggct 480
gagcacctgc tgcagacggc cgggacctgc ggggtcagcg acgacagcct gggcagcctc 540
ctgggacccc ggacggcagc cgtcttcagg cctcggcccg gggactctct gccatcccga 600
gagaccgcgt acgtggagct gtatgtggtc gtggacaatg cagagttcca gatgctgggg 660
agcgaagcag ccgtgcgtca tcgggtgctg gaggtggtga atcacgtgga caagctatat 720
cagaaactca acttcctgtg ggtcctggtg ggccctggaga tttggaatag tcaggacagg 780
ttccacgtca gccccgacct cagtgtcaca ctggagaacc tcctgacctg gcaggcacgg 840
caacggacac ggcggcacct gcatgacaac gtacagctca tcacgggtgt cgacttcacc 900
gggactactg tggggtttgc caggggtgtc gccatgtgct cccacagctc aggggctgtg 960
aaccaggacc acagcaagaa ccccggtggc gtggcctgca ccatggccca tgagatgggc 1020

```

```

cacaacctgg gcatggacca tgatgagaac gtccagggct gccgctgcca ggaacgcttc 1080
gaggccggcc gctgcatcat ggcaggcagc attggctcca gtttcccag gatgttcagt 1140
gactgcagcc aggcctacct ggagagcttt ttggagcggc cgcagtcggt gtgctcgcc 1200
aacgccccctg acctcagcca cctggtgggc ggccccgtgt gtgggaacct gtttgtggag 1260
cgtggggagc agtgcgactg cggccccccc gaggactgcc ggaaccgctg ctgcaactct 1320
accacctgcc agctggctga gggggcccag tgtgcgcacg gtacctgctg ccaggagtgc 1380
aagggtgaagc cggctgggtga gctgtgccgt cccaagaagg acatgtgtga cctcaggag 1440
ttctgtgacg gccggcaccg tgagtgcctg gaagacgcct tccaggagaa cggcacgccc 1500
tgctccgggg gctactgcta caacggggcc tgtcccacac tggcccagca gtgccaggcc 1560
ttctgggggg cagggtgggca ggctgccgag gagtccctgt tctcctatga catcctacca 1620
ggctgcaagg ccagcgagga cagggtgac atgtgtggcg ttctgcagtg caagggtggg 1680
cagcagcccc tggggcgctg catctgcata gtggatgtgt gccacgcgt caccacagag 1740
gatggcactg cgtatgaacc agtgcctgag ggcacccggt gtggaccaga gaaggtttgc 1800
tggaaggagc gttgccagga ctacacgtt tacagatcca gcaactgctc tgcccagtgc 1860
cacaaccatg ggggtgtgcaa ccacaagcag gagtgcact gccacgcggg ctgggccccg 1920
ccccactgcg cgaagctgct gactgaggtg cagcagcgt ccgggagcct ccccgctcct 1980
gtggtggtgg ttctggtgct cctggcagtt cctgtggtca ccctggcagg catcatcgct 2040
taccgcaaag ccggcagcg catctgagc aggaacgtgg ctcccaagac cacaatgggg 2100
cgctccaacc ccctgttcca ccaggctgcc agccgcgtgc cggccaaggg cggggctcca 2160
gccccatcca gggggcccca agagctgggt cccaccacc acccgggcca gcccgcccga 2220
caccggcct cctcggtggc tctgaagagg ccgccccctg ctccctcggg cactgtgtcc 2280
agcccaccct tcccagttcc tgtctacacc cggcaggcac caaagcaggt catcaagcca 2340
acgttcgcac cccagtgcc ccagtcaaa ccgggggctg gtgcggccaa ccctggtcca 2400
gctgagggtg ctgttggccc aaagggtgcc ctgaagcccc ccattccagag gaagcaagga 2460
gccggagctc ccacagcacc ctaggggggc acctgcgcct gtgtggaaat ttggagaagt 2520
tgccggcagag aagccatgcg ttccagcctt ccacgggtca gctagtgcg ctccagcccta 2580
gaccctgact ttgcaggctc agctgctggt ctaacctcag taatgcattt acctgagagg 2640
ctcctgctgt ccacgcctc agccaattcc ttctccccgc cttggccacg tgtagccccca 2700
gctgtctgca ggcaccaggc tgggatgagc tgtgtgcttg cgggtgcgtg tgtgtgtacg 2760
tgtctccagg tggccgctgg tctcccgtg tgttcaggag gccacatata cagccctcc 2820
cagccacacc tgccctgct ctggggcctg ctgagccggc tgccctgggc acccggttcc 2880
aggcagcaca gacgtggggc atccccagaa agactccatc ccaggaccag gttccccctc 2940
gtgctcttcg agagggtgtc agtgagcaga ctgcaccca agctcccgac tccagggtcc 3000
ctgatcttgg gcctgtttcc catgggattc aagagggaca gccccagctt tgtgtgtgtt 3060
taagcttagg aatgccccct atggaaaggg ctatgtggga gagtgcagta tcttgtctgg 3120
tttcttgag acctcagatg tgtgttcagc agggctgaaa gcttttattc ttaataatg 3180
agaaatgtat attttactaa taaattattg accgagttct gtagattctt gttaga 3236

```

<210> 60

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001109

<400> 60

ctttatggaa agggctatgt gggagagtca gctatcttgt ctggttttct tgagacctca 60

<210> 61

<211> 1449

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001124

<400> 61

```

ctggatagaa cagctcaagc cttgccactt cgggcttctc actgcagctg ggcttggact 60
tcggagtttt gccattgcca gtgggacgtc tgagactttc tcttcaagt acttggcaga 120
tcactctctt agcagggtct gcgttcgca gccgggatga agctggtttc cgtcgccctg 180

```

```

atgtacctgg  gttegetcgc  cttcctaggg  gctgacaccg  ctcggttgga  tgtcgcgtcg  240
gagtttccgaa  agaagtggaa  taagtgggct  ctgagtcgtg  ggaagaggga  actgcggatg  300
tccagcagct  accccaccgg  gctcgtcgac  gtgaaggccg  ggcctgccc  gacccttatt  360
cggccccagg  acatgaagg  tgctctcga  agccccgaag  acagcagtc  ggatgccgcc  420
cgcatccgag  tcaagcgcta  ccgccagagc  atgaacaact  tccagggcct  ccggagcttt  480
ggctgccgct  tcgggacgtg  cacgggtgcag  aagctggcac  accagatcta  ccagttcaca  540
gataaggaca  aggacaacgt  cgcccccagg  agcaagatca  gccccagg  ctacggccgc  600
cggcgccggc  gctccctgcc  cgaggccggc  cggggtcgga  ctctgggtgc  ttctaagcca  660
caagcacacg  gggctccagc  ccccccgagt  ggaagtgc  cccactttct  ttaggattta  720
ggcgcccatg  gtacaaggaa  tagtcgcgca  agcatccgc  tgggtgcctcc  cgggacgaag  780
gacttcccg  gcggtgtgg  gaccgggctc  tgacagccct  gcggagacc  tgagtccggg  840
aggcaccgct  cggcgcgag  ctctggcttt  gcaagggcc  ctccctctgg  gggcttcgct  900
tccttagcct  tgctcaggtg  caagtgcctc  agggggcg  gtgcagaaga  atccgagtg  960
ttgccaggct  taaggagagg  agaaactgag  aaatgaatgc  tgagacccc  ggagcaggg  1020
tctgagccac  agcctgtctc  gccacaaac  tgatttctca  cggcgtgtca  cccaccagg  1080
gcgcaagcct  cactattact  tgaactttcc  aaaacctaaa  gaggaaaagt  gcaatgcgtg  1140
ttgtacatac  agaggtaact  atcaatattt  aagtttgtg  ctgtcaagat  tttttttgta  1200
acttcaaata  tagagatatt  tttgtacgtt  atatattgta  ttaagggc  tttaaaagca  1260
atttatattgt  cctccctat  ttttaagacgt  gaatgtctca  gcgaggtgta  aagttgttcg  1320
ccgctggaa  tgtgagtg  tttgtgtgca  tgaaagagaa  agactgatta  cctcctgtgt  1380
ggaagaagga  aacaccgagt  ctctgtataa  tctatttaca  taaaatgggt  gatatgcgaa  1440
cagcaaacc  1449

```

<210> 62
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001124

<400> 62
 gaaggaaaca ccgagtctct gtataatcta tttacataaa atgggtgata tcggaacagc 60

<210> 63
 <211> 1619
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001168

```

<400> 63
ccgccagatt  tgaatcgccg  gaccctgttg  cagaggtggc  ggccggcgga  tgggtgcccc  60
gacgttgccc  cctgcctggc  agccctttct  caaggaccac  cgcatctcta  cattcaagaa  120
ctggcccttc  ttggagggtc  gcgcctgcac  cccggagcgg  atggccgagg  ctggcttcac  180
ccactgcccc  actgagaacg  agccagactt  ggcccagtg  ttcttctgct  tcaaggagct  240
ggaaggctgg  gagccagatg  acgaccccat  agaggaacat  aaaaagcatt  cgtccggttg  300
cgctttcctt  tctgtcaaga  agcagtttga  agaattaacc  cttggtgaat  ttttgaaact  360
ggacagagaa  agagccaaga  acaaaattgc  aaaggaaacc  aacaataaga  agaaagaatt  420
tgaggaaact  gcgaagaaag  tgcgcctg  catcgagcag  ctggctgcca  tggattgagg  480
cctctggccg  gagctgcctg  gtcccagagt  ggctgcacca  cttccagggt  ttattccctg  540
gtgccaccag  ccttcctgtg  ggccccttag  caatgtotta  ggaaaggaga  tcaacatttt  600
caaattagat  gtttcaactg  tgctcctg  ctggtaacag  tggctgcttc  tctctctctc  660
tgctgtgca  gcgggtgctg  ctggtaacag  tggctgcttc  tctctctctc  tctctttttt  720
gggggctcat  ttttgcctgt  ttgattcccg  ggcttaccag  gtgagaagtg  agggaggaag  780
aaggcagtg  cccttttgc  agagctgaca  gctttgttcg  cgtgggcaga  gccttcacac  840
gtgaatgtgt  ctggacctca  tgttgttgag  gctgtcacag  tcttgagtg  ggacttggca  900
ggtgcctgtt  gaatctgagc  tgcaggttcc  ttatctgtca  cacctgtgcc  tctcagagg  960
acagtttttt  tgttgttgtg  tttttttgtt  tttttttttt  gtagatgca  tgacttgtgt  1020
gtgatgagag  aatggagaca  gagtccctgg  ctctctact  gtttaacaac  atggctttct  1080

```



```

tattttgttt gaattgttaa ttcacagaat agcacaaact acaattaaaa ctaagcacaa 1140
agccattcta agtcattggg gaaacggggg gaacttcagg tggatgagga gacagaatag 1200
agtgatagga agcgtctggc agatactcct tttgccactg ctgtgtgatt agacaggccc 1260
agtgagccgc ggggcacatg ctggccgctc ctccctcaga aaaaggcagt ggcctaaatc 1320
ctttttaaag gacttggctc gatgctgtgg gggactggct gggctgctgc aggccgtgtg 1380
tctgtcagcc caaccttcac atctgtcacg ttctccacac gggggagaga cgcagtccgc 1440
ccagggtcccc gctttctttg gaggcagcag ctcccgagg gctgaagtct ggcgtaagat 1500
gatggatttg attcgccctc ctccctgtca tagagctgca ggggtggattg ttacagcttc 1560
gctggaaacc tctggaggtc atctcggtg ttcttgagaa ataaaaagcc tgtcatttc 1619

```

```

<210> 64
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001168

```

```

<400> 64
ttcacagaat agcacaaact acaattaaaa ctaagcacaa agccattcta agtcattggg 60

```

```

<210> 65
<211> 1552
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001216

```

```

<400> 65
gccgtacac accgtgtgct gggacacccc acagtcagcc gcatggctcc cctgtgcccc 60
agcccttggc tccctctgtt gatcccgccc cctgtctcag gcctcactgt gcaactgctg 120
ctgtcactgc tgcctctgat gcctgtccat cccagagggt tgccccggat gcaggaggat 180
tcccccttgg gaggaggctc ttctggggaa gatgaccac tgggagagga ggatctgccc 240
agtgaagagg attcaccacag agaggaggat ccaccggag aggaggatct acctggagag 300
gaggatctac ctggagagga ggatctacct gaagttaagc ctaaatcaga agaagagggc 360
tccctgaagt tagaggatct acctactgtt gaggctcctg gagatcctca agaaccctag 420
aataatgccc acagggacaa agaaggggat gaccagagtc attggcgcta tggaggcgac 480
ccgccttggc cccgggtgtc cccagcctgc gggggccgct tccagtcccc ggtggatatc 540
cgccccagc tcgcccctt ctgcccggcc ctgcgcccc tggaaactct gggcttccag 600
ctcccgccgc tcccagaact gcgcctgcgc aacaatggcc acagtgtgca actgaccctg 660
cctctggggc tagagatggc tctgggtccc gggcgggagt accgggctct gcagctgcat 720
ctgcaactgg gggctgcagg tcgtccgggc tcggagcaca ctgtggaagg ccaccgtttc 780
cctgccgaga tccactggtt tcacctcagc accgcctttg ccagagttga cgaggccttg 840
gggcgccccg gaggcctggc cgtgttggcc gcctttctgg aggaggggcc ggaagaaaac 900
agtgcctatg agcagttgct gtctcgcttg gaagaaatcg ctgaggaagg ctacagagat 960
cagggtcccag gactggacat atctgcactc ctgcccctctg acttcagccg ctacttccaa 1020
tatgaggggt ctctgactac accgcctgtt gccaggggtg tcatctggac tgtgtttaac 1080
cagacagtga tgctgagtgc taagcagctc cacaccctct ctgacacct gtggggacct 1140
ggtgactctc ggctacagct gaacttccga gcgacgcagc ctttgaatgg gcgagtgtt 1200
gaggcctcct tccctgctgg agtggacagc agtcctcggg ctgctgagcc agtcagctg 1260
aattcctgcc tggctgctgg tgacatccta gccctgggtt ttggcctcct tttgtctgtc 1320
accagcgtcg cgttccttgt gcagatgaga aggcagcaca gaaggggaac caaagggggg 1380
gtgagctacc gccagcaga ggtagccgag actggagcct agaggctgga tcttgagaaa 1440
tgtgagaagc cagccagagg catctgagg ggagccggt actgtcctgt cctgctcatt 1500
atgccacttc cttttaactg ccaagaaatt ttttaaaata aatatttata at 1552

```

```

<210> 66
<211> 60
<212> DNA
<213> Homo sapiens

```

<300>

<308> NM_001216

<400> 66

tcctgtcctg ctcattatgc cacttccttt taactgccaa gaaatttttt aaaataaata 60

<210> 67

<211> 2653

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001254

<400> 67

gagcgcggct ggagtttgc tctgccgctg tgcagtttgc tcaggggctt gtgggtggtga 60
 gtccgagagg ctgcgtgtga gagacgtgag aaggatcctg cactgaggag gtggaaagaa 120
 gaggattgct cgaggaggcc tggggtctgt gaggcagcgg agctgggtga aggctgcggg 180
 ttccggcgag gcctgagctg tgcgtgcgtc atgcctcaaa cccgatccca ggcacaggct 240
 acaatcagtt ttccaaaaag gaagctgtct cgggcattga acaaagctaa aaactccagt 300
 gatgccaaac tagaaccaac aaatgtccaa accgtaacct gttctcctcg tgtaaaagcc 360
 ctgcctctca gcccaggaa acgtctgggc gatgacaacc tatgcaacac tccccattta 420
 cctccttggt ctccaccaa gcaaggcaag aaagagaatg gtccccctca ctcacatata 480
 cttaagggac gaagattggt atttgacaat cagctgacaa ttaagtctcc tagcaaaaga 540
 gaactagcca aagttcacca aaacaaaata ctttcttcag ttagaaaaag tcaagagatc 600
 acaacaaatt ctgagcagag atgtccactg aagaaagaat ctgcatgtgt gagactatcc 660
 aagcaagaag gcacttgcta ccagcaagca aagctgggtc tgaacacagc tgtcccagat 720
 cggctgcctg ccagggaag ggagatggat gtcacacagga atttcttgag ggaacacatc 780
 tgtgggaaaa aagctggaag cctttacctt tctggtgctc ctggaactgg aaaaactgcc 840
 tgcttaagcc ggattctgca agacctcaag aagggaactga aaggctttaa aactatcatg 900
 ctgaattgca tgccttgag gactgccag gctgtattcc cagctattgc tcaggagatt 960
 tgtcaggaag aggtatccag gccagctggg aaggacatga tgaggaaatt ggaaaaacat 1020
 atgactgcag agaagggcc catgattgtg ttggtattgg acgagatgga tcaactggac 1080
 agcaaaggcc aggatgtatt gtacacgcta tttgaatggc catggctaag caattctcac 1140
 ttggtgctga ttggtattgc taataccctg gatctcacag atagaattct acctaggctt 1200
 caagctagag aaaaatgtaa gccacagctg ttgaaactcc cacttatac cagaaatcag 1260
 atagtcacta ttttgcaaga tgcacttaat caggatctca gagatcaggt tctggacaat 1320
 gctgcagttc aattctgtgc ccgcaaagtc tctgctgttt caggagatgt tcgcaaagca 1380
 ctggatgttt gcaggagagc tattgaaatt gtagagtcag atgtcaaaag ccagactatt 1440
 ctcaaaccac tgcctgaatg taaatcacct tctgagcctc tgattcccaa gaggggttgg 1500
 cttattcaca tatcccaagt catctcagaa gttgatggta acaggatgac cttgagccaa 1560
 gaaggagcac aagattcctt cctcttcag cagaagatct tggtttgctc tttgatgctc 1620
 ttgatcaggc agttgaaat caaagaggtc actctgggga agttatatga agcctacagt 1680
 aaagtctgtc gcaaacagca ggtggcggct gtggaccagt cagagtgttt gtcactttca 1740
 gggctcttgg aagccagggg cattttagga ttaaagagaa acaaggaaac ccgtttgaca 1800
 aaggtgtttt tcaagattga agagaaagaa atagaacatg ctctgaaaga taaagcttta 1860
 attggaaata tcttagctac tggattgcct taaattcttc tcttacaccc cacccgaaag 1920

tattcagctg gcatttagag agctacagtc ttcatTTTTag tgcttttacac attcgggctt 1980
 gaaaacaaat atgacctttt ttacttgaag ccaatgaatt ttaattctata gattctttta 2040
 tattagcaca gaataatatc tttgggtctt actatTTTTa ccataaaaag tgaccaggta 2100
 gacctttttt aattacattc actacttcta ccacttggtg atctctagcc aatgtgcttg 2160
 caagtgtaca gatctgtgta gaggaatgtg tgtatatTTa cctcttctgt tgcTcaaaCa 2220
 tgagtgggta tttttttgtt tgtttttttt gttgttgggt tttttgaggc gcgtctcacc 2280
 ctgttgccca ggctggagtg caatggcgcg ttctctgctc actacagcac ccgcttccca 2340
 gggtgaagtg attctcttgc ctacagcctc cgagtagctg ggattacagg tgcaccacc 2400
 cgcgcccagc taatttttta atttttagta gagacagggt tttaccatgt tggccaggct 2460
 ggtcttgaac tcctgaccct caagtgatct gccaccttg gcctccctaa gtgctgggat 2520
 tataggcgtg agccaccatg ctacagccatt aaggtatttt gttaagaact ttaagtttag 2580
 ggtaagaaga atgaaaatga tccagaaaaa tgcaagcaag tccacatgga gatttggagg 2640

acactggtta aag 2653

<210> 68
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001254

<400> 68
 caaggaaacc cgtttgacaa aggtgttttt caagattgaa gagaaagaaa tagaacatgc 60

<210> 69
 <211> 627
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001323

<400> 69
 gcggccgcaa gctcggcact cacggctctg agggctccga cggcactgac ggccatggcg 60
 cgttcgaacc tcccgcctggc gctgggcctg gccctggtcg cattctgcct cctggcgctg 120
 ccacgcgacg cccgggcccc gccgcaggag cgcattggtcg gagaactccg ggacctgtcg 180
 cccgacgacc cgcaggtgca gaaggcggcg caggcggccg tggccagcta caacatgggc 240
 agcaacagca tctactactt ccgagacacg cacatcatca aggcgcagag ccagctggtg 300
 gccggcatca agtacttccct gacgatggag atggggagca cagactgccg caagaccagg 360
 gtcactggag accacgtcga cctcaccact tgccccctgg cagcaggggc gcagcaggag 420
 aagctgcgct gtgactttga ggtccttggt gttccctggc agaactcctc tcagctccta 480
 aagcacaact gtgtgcagat gtgataagtc cccgagggcg aaggccattg ggtttggggc 540
 catggtggag ggcacttcag gtccgtgggc cgtatctgtc acaataaatg gccagtgctg 600
 cttcttgcaa aaaaaaaaaa aaaaaaa 627

<210> 70
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001323

<400> 70
 atcaagtact tcctgacgat ggagatgggg agcacagact gccgcaagac cagggtcact 60

<210> 71
 <211> 1812
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001428

<400> 71
 tagctaggca ggaagtccgc gcgggcggcg cggacagtat ctgtgggtac ccggagcacg 60
 gagatctcgc cggtctttacg ttcacctcgg tgtctgcagc accctccgct tcctctccta 120
 ggcgacgaga cccagtggct agaagttcac catgtctatt ctcaagatcc atgccaggga 180
 gatctttgac tctcgcgga atcccactgt tgagggtgat ctcttcacct caaaagggtct 240
 cttcagagct gctgtgccca gtggtgcttc aactgggtatc tatgaggccc tagagctccg 300
 ggacaatgat aagactcgct atatggggaa ggggtgtctca aaggctgttg agcacatcaa 360
 taaaactatt gcgcctgccc tggttagcaa gaaactgaac gtcacagaac aagagaagat 420

```

tgacaaactg atgatcgaga tggatggaac agaaaataaa tctaagtttg gtgcgaacgc 480
cattctgggg gtgtcccttg ccgtctgcaa agctgggtgcc gttgagaagg ggggtccccc 540
gtaccgccac atcgctgact tggctggcaa ctctgaagtc atcctgccag tcccggcggt 600
caatgtcatc aatggcggtt ctcatgctgg caacaagctg gccatgcagg agttcatgat 660
cctcccagtc ggtgcagcaa acttcagggg agccatgcgc attggagcag aggtttacca 720
caacctgaag aatgtcatca aggagaaata tgggaaagat gccaccaatg tgggggatga 780
aggcggtttt gctcccaaca tcctggagaa taaagaaggc ctggagctgc tgaagactgc 840
tattgggaaa gctgggtaca ctgataagggt ggtcatcggc atggacgtag cggcctccga 900
gttcttcagg tctgggaagt atgacctgga cttcaagtct cccgatgacc ccagcaggta 960
catctcgctt gaccagctgg ctgacctgta caagtccttc atcaaggact acccagtggg 1020
gtctatcgaa gatccctttg accaggatga ctggggagct tggcagaagt tcacagccag 1080
tgcaggaatc caggtagtgg gggatgatct cacagtgacc aacccaaaga ggatcgccaa 1140
ggccgtgaac gagaagtctt gcaactgcct cctgctcaaa gtcaaccaga ttggctccgt 1200
gaccgagctt cttcaggcgt gcaagctggc ccaggccaat ggttggggcg tcatggtgtc 1260
tcatcgcttg ggggagactg aagatacctt catcgctgac ctggttgtgg ggctgtgcac 1320
tgggcagatc aagactggtg ccccttgccg atctgagcgc ttggccaagt acaaccagct 1380
cctcagaatt gaagaggagc tgggcagcaa ggctaagttt gccggcagga acttcagaaa 1440
ccccttggcc aagtaagctg tgggcaggca agcccttcgg tcacctgttg gctacacaga 1500
cccctccctt cgtgtcagct caggcagctc gaggcccccg accaactt gcaggggtcc 1560
ctgctagtta gcgccccacc gcgctggagt tcgtaccgct tccttagaac ttctacagaa 1620
gccaaagctc ctggagccct gttggcagct ctagctttgc agtcgtgtaa ttggcccaag 1680
tcattgtttt tctcgctca ctttccacca agtgtctaga gtcatgtgag cctcgtgtca 1740
tctccggggt ggccacaggc tagatccccg gtggttttgt gctcaaaata aaaagcctca 1800
gtgacccatg ag 1812

```

<210> 72

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001428

<400> 72

```

agctctagct tttgcagtcg tgtaatgggc ccaagtcatt gtttttctcg cctcactttc 60

```

<210> 73

<211> 8368

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001456

<400> 73

```

gcgatccggg cgccaccccg cggtcacatg tcaccggctg ctctcaggaa cagcagcgca 60
acctctgctc cctgcctcgc ctcccgcgcc cctaggtgcc tgcgacttta attaaagggc 120
cgtccctctg ccgaggtctg agcaocgccc ccccggttcc tcgcgcctca aaatgagtag 180
ctcccaactc cgggcggggc agagcgcagc aggcgcggct ccgggcggcg gcgtcgacac 240
gcgggacgcc gagatgcccg ccacogagaa ggacctggcg gaggacgcgc cgtggaagaa 300
gatccagcag aacactttca cgcgctgggt caacgagcac ctgaagtgcg tgagcaagcg 360
catcgccaac ctgcagacgg acctgagcga cgggctgcgg cttatcgcgc tgttggaggt 420

```

```

gctcagccag aagaagatgc accgcaagca caaccagcgg cccactttcc gccaaatgca 480
gcttgagaac gtgtcggttg cgctcgagtt cctggaccgc gagagcatca aactggtgtc 540
catcgacagc aaggccatcg tggacgggaa cctgaagctg atcctgggcc tcatctggac 600
cctgatcctg cactactcca tctccatgcc catgtgggac gaggaggagg atgaggaggc 660
caagaagcag acccccagc agaggctcct gggctggatc cagaacaagc tgccgcagct 720
gcccatcacc aacttcagcc gggactggca gagcggccgg gccctgggcg ccctggtgga 780
cagctgtgcc ccgggcctgt gtctctgact ggactcttgg gacgccagca agcccgttac 840

```

caatgcgcga	gaggccatgc	agcaggcgga	tgactggctg	ggcatcccc	aggtgatcac	900
ccccgaggag	attgtggacc	ccaacgtgga	cgagcactct	gtcatgacct	acctgtccca	960
gttccccaa	gccaagctga	agccaggggc	tcccttgccg	cccaaactga	acccgaagaa	1020
agcccggtgc	tacgggccag	gcacgcagcc	cacaggcaac	atgggtgaaga	agcgggcaga	1080
gttccactgtg	gagaccagaa	gtgctggcca	gggagagggtg	ctgggtgtacg	tggagaccc	1140
ggccggacac	caggaggagg	caaaagtgc	cgccaataac	gacaagaacc	gcaccttctc	1200
cgtctggtac	gtccccgagg	tgacggggac	tcataagggtt	actgtgctct	ttgctggcca	1260
gcacatcgcc	aagagcccct	tcgaggtgta	cgtggataag	tcacagggtg	acgccagcaa	1320
agtgcagacc	caagggtccc	gcctggagcc	cagtggcaac	atcgccaaca	agaccaccta	1380
ctttgagatc	tttacggcag	gagctggcac	gggcgagggtc	gaggttggtga	tccaggacc	1440
catgggacag	aagggccacg	tagagcctca	gttgaggcc	cggggcgaca	gcacataccg	1500
ctgcagctac	cagcccacca	tggaggcggt	ccacaccgtg	cacgtcacgt	ttgccggcgt	1560
gcccatccct	cgcagcccct	acactgtcac	tgttggccaa	gcctgtaacc	cgagtgcctg	1620
ccgggcgggtt	ggccggggcc	tccagcccaa	gggtgtgcgg	gtgaaggaga	cagctgactt	1680
caaggtgtac	acaaagggcg	ctggcagtg	ggagctgaag	gtcaccgtga	agggccccaa	1740
gggagaggag	cgcgtgaagc	agaaggacct	gggggatggc	gtgtatggct	tcgagtatta	1800
cccctatgggt	cctggaacct	atatcgctac	catcacgtgg	ggtgggtcaga	acatcgggcg	1860
cagtcaccttc	gaagtgaagg	tgggcaccga	gtgtggcaat	cagaaggtag	gggcctgggg	1920
ccttgggctg	gaggggcgcg	tcgttggcaa	gtcagcagac	tttgtggtgg	aggctatcgg	1980
ggacgacgtg	ggcacgctgg	gcttctcggt	ggaagggcca	tcgcaggcta	agatcgaatg	2040
tgacgacaag	ggcgacggct	cctgtgatgt	gcgctactgg	ccgcaggagg	ctggcgagta	2100
tgccgttcac	gtgctgtgca	acagcgaaga	catccgcctc	agcccttca	tggctgacat	2160
ccgtgacg	ccccaggact	tccaccaga	cagggtgaag	gcacgtgggc	ctggattgga	2220
gaagacaggt	gtggccgtca	acaagccagc	agagttcaca	gtggatgcca	agcacgggtg	2280
caaggcccca	cttcgggtcc	aagtccagga	caatgaaggc	tgcctgtgg	aggcgttgg	2340
caaggacaac	ggcaatggca	cttacagctg	ctcctacgtg	cccaggaagc	cggtgaagca	2400
cacagccatg	gtgtcctggg	gaggcgctag	catccccaac	agcccttca	gggtgaatgt	2460
gggagctggc	agccacccca	acaagggtca	agtatacggc	cccggagtag	ccaagacagg	2520
gctcaaggcc	cacgagccca	cctacttcac	tgtggactgc	gccgaggctg	gccaggggga	2580
cgtcagcatc	ggcatcaagt	gtgcccctgg	agtggtaggc	cccgcgaag	ctgacatcga	2640
cttcgacatc	atccgcaatg	acaatgacac	cttcacggtc	aagtagacgc	cccggggggc	2700
tggcagctac	accattatgg	tctctttg	tgaccaggcc	acgcccacca	gccccatccg	2760
agtcaagggtg	gagccctctc	atgacgccag	taagggtgaag	gccgagggcc	ctggcctcag	2820
tcgcactgg	gtcgagcttg	gcaagcccac	ccacttcaca	gtaaattgcca	aagctgctgg	2880
caaaggcaag	ctggacgtcc	agttctcagg	actcaccaag	ggggatgcag	tgcgagatgt	2940
ggacatcatc	gaccaccatg	acaacaccta	cacagtcagg	tacacgcctg	tccagcagg	3000
tccagttaggc	gtcaatgtca	cttatggagg	ggatcccatc	cctaagagcc	ctttctcagt	3060
ggcagtatct	ccaagcctgg	acctcagcaa	gatcaagggtg	tctggcctgg	gagagaagg	3120
ggacgttggc	aaagaccagg	agttcacagt	caaatacaag	ggtgctgggtg	gtcaaggcaa	3180
agtggcatcc	aagattgtgg	gcccctcggg	tgcagcggtg	ccctgcaagg	tggagccagg	3240
cctgggggct	gacaacagt	tgggtgcgctt	cctgcccctg	gaggaagggc	cctatgaggt	3300
ggaggtgacc	tatgacggcg	tgcccgtg	tggcagcccc	tttctctctg	aagctgtggc	3360
ccccccaag	cctagcaagg	tgaaggcggt	tgggccgggg	ctgcagggag	gcagtcggg	3420
ctcccccgcc	cgcttcacca	tcgacaccaa	gggcgcgggc	acagggtggcc	tgggcctgac	3480
ggtggaggggc	ccctgtgagg	cgcagctcga	gtgcttggac	aatggggatg	gcacatgttc	3540
cgtgtcctac	gtgcccaccg	agcccgggga	ctacaacatc	aacatcctct	tcgctgacac	3600
ccacatccct	ggctcccat	tcaaggccca	cgtgggtccc	tgttttgacg	catccaaagt	3660
caagtgtc	ggccccgggc	tggagcgggc	caccgctggg	gaggtggg	aattccaagt	3720
ggactgtctg	agcgcgggga	gcgcggagct	gaccattgag	atctgctcgg	aggcggggct	3780
tccggccgag	gtgtacatcc	aggaccacgg	tgatggcagc	cacaccatta	cctacattcc	3840
cctctgcccc	ggggcctaca	ccgtcaccat	caagtagggc	ggccagcccg	tgcccactt	3900
ccccagcaag	ctgcaggtgg	aacctgcggt	ggacacttcc	ggtgtccagt	gctatggg	3960
tggatttgag	ggccaggggtg	tcttccgtga	ggccaccaact	gagttcagtg	tggacgccc	4020
ggctctgaca	cagaccggag	ggccgcacgt	caaggcccgt	gtggccaacc	cctcaggcaa	4080
cctgacggag	acctacgttc	aggaccgtgg	cgatggcatg	tacaaagtgg	agtacacgcc	4140
ttacgaggag	ggactgcact	ccgtggacgt	gacctatgac	ggcagtcocg	tgcccagcag	4200
ccccttcag	gtgcccgtga	ccgagggctg	cgacccctcc	cgggtgcgtg	tccacggggc	4260
aggcatccaa	agtggcacca	ccaacaagcc	caacaagttc	actgtggaga	ccaggggagc	4320
tggcacgggc	ggcctggg	tggctgtaga	gggcccctcc	gaggccaaga	tgtcctgcat	4380
ggataacaag	gacggcagct	gctcggtcga	gtacatccct	tatgaggctg	gcacctacag	4440
cctcaacgtc	acctatgggt	gccatcaagt	gccaggcagt	cctttcaagg	tccctgtgca	4500

tgatgtgaca	gatgcgtcca	aggtcaagtg	ctctggggccc	ggcctgagcc	caggcatggt	4560
tcgtgccaac	ctccctcagt	ccttccagggt	ggacacaagc	aaggctggtg	tggccccatt	4620
gcaggtcaaa	gtgcaagggc	ccaaaggcct	ggtggagcca	gtggacgtgg	tagacaacgc	4680
tgatggcacc	cagaccgtca	attatgtgcc	cagccgagaa	gggccctaca	gcattctcagt	4740
actgtatgga	gatgaagagg	taccccgagg	ccccttcaag	gtcaagggtg	tgccctactca	4800
tgatgccagc	aagggtgaagg	ccagtggccc	cgggctcaac	accactggcg	tgccctgccag	4860
cctgcccgtg	gagttcacca	tcgatgcaaa	ggacgcgggg	gagggcctgc	tggctgtcca	4920
gatcacggat	cccgaaggca	agccgaagaa	gacacacatc	caagacaacc	atgacggcac	4980
gtatacagtg	gcctaogtgc	cagacgtgac	aggtcgctac	accatcctca	tcaagtacgg	5040
tggtgacgag	atcccttctt	cccgtaccg	cgtgcgtgcc	gtgcccaccg	gggacggccag	5100
caagtgcact	gtcacagtgt	caatcggagg	tcacgggcta	ggtgctggca	tcggcccacc	5160
cattcagatt	ggggaggaga	cgggtgatcac	tgtggacact	aaggcggcag	gcaaaggcaa	5220
agtgcagtgc	accgtgtgca	cgcctgatgg	ctcagaggtg	gatgtggacg	tgggtggagaa	5280
tgaggacggc	acttttcgaca	tcttctacac	ggccccccag	ccggggcaaat	acgtcatctg	5340
tgtgcgcttt	ggtggcgagc	acgtgcccac	cagccccttc	caagtgcagg	ctctggctgg	5400
ggaccagccc	tcggtgcagc	cccctctacg	gtctcagcag	ctggccccac	agtacacctc	5460
cgcccagggc	ggccagcaga	cttggggccc	ggagaggccc	ctggtgggtg	tcaatgggct	5520
ggatgtgacc	agcctgaggc	cctttgacct	tgtcatcccc	ttcaccatca	agaaggcgca	5580
gatcacaggg	gaggttcgga	tgccctcagg	caaggtggcg	cagcccacca	tcactgacaa	5640
caaagacggc	accgtgaccg	tgcggtatgc	accagcgag	gctggcctgc	acgagatgga	5700
catccgctat	gacaacatgc	acatcccagg	aagccccttg	cagttctatg	tggattacgt	5760
caactgtggc	catgtcactg	cctatggggc	tggcctcacc	catggagtag	tgaacaagcc	5820
tgccaccttc	accgtcaaca	ccaaggatgc	aggagagggg	ggcctgtctc	tggccattga	5880
gggcccgtcc	aaagcagaaa	tcagctgcac	tgacaaccag	gatgggacat	gcagcgtgtc	5940
ctacctgcct	gtgctgccgg	gggactacag	cattctagtc	aagtacaatg	aacagcagct	6000
cccaggcagc	cccttcactg	ctcggtgcac	aggtgacgac	tccatgcgta	tgtcccacct	6060
aaaggctggc	tctgctgccc	acatccccat	caacatctca	gagacggatc	tcagcctgct	6120
gacggccact	gtggtcccg	cctcgggccg	ggaggagccc	tgtttgctga	agcggctgctg	6180
taatggccac	gtggggattt	cattcgtgcc	caaggagacg	ggggagcacc	tggtgcatgt	6240
gaagaaaaat	ggccagcagc	tggccagcag	ccccatccc	gtggtgatca	gccagtcgga	6300
aattggggat	gccagtcgtg	ttcggtgtct	tggctcaggc	cttcacgaag	gccacacctt	6360
tgagcctgca	gagtttatca	ttgatacccg	cgatgcaggc	tatggtgggc	tcagcctgtc	6420
cattgagggc	cccagcaagg	tggacatcaa	cacagaggac	ctggaggacg	ggacgtgcag	6480
ggtcacctac	tgccccacag	agccaggcaa	ctacatcatc	aacatcaagt	ttgccgacca	6540
gcacgtgcct	ggcagccctt	tctctgtgaa	ggtgacaggc	gagggccggg	tgaagagagag	6600
gcatcccgcc	aggcgtcggg	ctccttcagt	ggccaacgtt	ggtagtcatt	gtgacctcag	6660
cctgaaaatc	cctgaaatta	gcattccagga	tatgacagcc	caggtgacca	gcccacgggg	6720
caagacccat	gagggccgaga	tcgtggaagg	ggagaaccac	acctactgca	tccgcttttgt	6780
tcccgtctgag	atgggcacac	acacagtcag	cgtcaagtac	aagggccagc	acgtgccttg	6840
gagcccttcc	cagttcaccc	tggggccccc	aggggaaggg	ggagcccaca	aggtccgagc	6900
tggggggccct	ggcctggaga	gagctgaagc	tggagtgcc	gccgaattca	gtatctggac	6960
ccgggaagct	ggtgctggag	gcctggccat	tgtgtctgag	ggcccagca	aggctgagat	7020
ctctttttgag	gaccgcaagg	acggctcctg	tgtgtggct	tatgtggtcc	aggagccagg	7080
tgactacgaa	gtctcagtca	agttcaacga	ggaacacatt	cccagacagc	ccttcgtggt	7140
gcctgtggct	tctccgtctg	gcgacgccc	ccgcctcact	gtttctagcc	ttcaggagtc	7200
agggctaaag	gtcaaccagc	cagcctcttt	tgcatgcagc	ctgaacgggg	ccaagggggc	7260
gatcgatgcc	aagggtgcaca	gcccctcagg	agccctggag	gagtgctatg	tcacagaaat	7320
tgaccaagat	aagtatgctg	tgcgcttcag	cctcggggag	aatggcggtt	acctgattga	7380
cgtcaagttc	aacgggtaccc	acatccctgg	aagcccttc	aagatccgag	ttggggagcc	7440
tgggcatgga	ggggacccag	gcttgggtgtc	tgcttacgga	gcaggctctg	aaggcgggtg	7500
cacagggaac	ccagctgagt	tcgtcgtgaa	cacagcaaat	gcgggagctg	gtgccctgtc	7560
ggtgaccatt	gacggccctt	ccaagggtgaa	gatggattgc	caggagtgc	ctgagggcta	7620
ccgcgtcacc	tataccccca	tggcacctgg	cagctacctc	atctccatca	agtacggcgg	7680
ccctaccac	attgggggca	gccccttcaa	ggccaaagtc	acaggccccc	gtctcgtcag	7740
caaccacagc	ctccacgaga	catcatcagt	gtttgtagac	tctctgacca	aggccacctg	7800
tgccccccag	catggggccc	cgggtcctgg	cgtgctgac	gccagcaagg	tgggtggcaa	7860
gggcctgggg	ctgagcaagg	cctacgtagg	ccagaagagc	agcttcacag	tagactgcag	7920
caaagcaggc	aacaacatgc	tgctggtggg	ggttcatggc	ccaaggaccc	cctgogagga	7980
gatcctgggtg	aagcacgtgg	gcagccggct	ctacagcgtg	tcctacctgc	tcaaggacaa	8040
gggggagtag	acactggtgg	tcaaatgggg	gcacgagcac	atcccaggca	gcccctaccg	8100
cgttggtggtg	ccctgagtct	ggggcccggtg	ccagccggca	gcccccaagc	ctgccccgct	8160

acccaagcag ccccgccctc tccccctcaa ccccggccca ggccgccttg gccgcccgc 8220
 tgtcactgca gctgcccctg cctgtgccc tgctgcgctc acctgcctcc ccagccagcc 8280
 gctgacctct cggctttcac ttgggcagag ggagccattt ggtggcgctg cttgtcttct 8340
 ttggttcttg gagggtgag ggatgggg 8368

<210> 74
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001456

<400> 74
 tgacctctcg gctttcactt gggcagaggg agccatttgg tggcgctgct tgtcttcttt 60

<210> 75
 <211> 1642
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001548

<400> 75
 ccagatctca gaggagcctg gctaagcaaa accctgcaga acggctgcct aatttacagc 60
 aaccatgagt acaaatgggtg atgatcatca ggtcaaggat agtctggagc aattgagatg 120
 tacttttaca tgggagttat ccattgatga cgatgaaatg cctgatttag aaaacagagt 180
 cttggatcag attgaattcc tagacaccaa atacagtgtg ggaatacaca acctactagc 240
 ctatgtgaaa cacctgaaaag gccagaatga ggaagccctg aagagcttaa aagaagctga 300
 aaacttaatg caggaagaac atgacaacca agcaaatgtg aggagtctgg tgacctgggg 360
 caactttgcc tggatgtatt accacatggg cagactggca gaagcccaga cttacctgga 420
 caagggtggag aacatttgca agaagctttc aaatcccttc cgctatagaa tggagtgtcc 480
 agaaatagac tgtgaggaag gatgggcctt gctgaagtgt ggaggaaaga attatgaacg 540
 ggccaaggcc tgctttgaaa aggtgcttga agtggaccct gaaaaccctg aatccagcgc 600
 tgggtatgcg atctctgcct atgcctgga tggctttaaa ttagccacaa aaaatcacia 660
 gccattttct ttgcttcccc taaggcaggc tgtccgctta aatccagaca atggatatat 720
 taaggttctc cttgccctga agcttcagga tgaaggacag gaagctgaag gagaaaagta 780
 cattgaagaa gctctagcca acatgtcctc acagacctat gtctttcgat atgcagccaa 840
 gttttaccga agaaaaggct ctgtggataa agctcttgag ttattaaaaa aggccttgca 900
 ggaaacaccc acttctgtct tactgcatca ccagataggg ctttgttaca aggcacaaat 960
 gatccaaatc aaggaggcta caaaagggca gcctagaggg cagaacagag aaaagctaga 1020
 caaaatgata agatcagcca tatttcattt tgaatctgca gtggaaaaaa agcccacatt 1080
 tgagggtggct catctagacc tggcaagaat gtatatagaa gcaggcaatc acagaaaagc 1140
 tgaagagaat tttcaaaaat tgttatgcat gaaaccagtg gtagaagaaa caatgcaaga 1200
 catacatttc tactatgggc ggtttcagga atttcaaaaag aaatctgacg tcaatgcaat 1260
 tatccattat ttaaaagcta taaaaataga acaggcatca ttaacaaggg ataaaagtat 1320
 caattctttg aagaaattgg ttttaaggaa acttcggaga aaggcattag atctggaaag 1380
 cttgagcctc cttgggttcg tctataaatt ggaaggaaat atgaatgaag ccctggagta 1440
 ctatgagcgg gccctgagac tggctgctga ctttgagaac tctgtgagac aaggctcctta 1500
 ggcacccaga tatcagccac ttccacattt catttcattt tatgctaaca ttactaatc 1560
 atctttcttg cttactgttt tcagaaacat tataattcac tgtaatgatg taattcttga 1620
 ataataaatc tgacaaaata tt 1642

<210> 76
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001548

<400> 76

gtatcaattc tttgaagaaa ttggtttttaa ggaaacttcg gagaaaggca ttagatctgg 60

<210> 77

<211> 3344

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001605

<400> 77

```

ggtacagctg cgcgtctgcg ggaataggtg cagcggggccc ttggcggggg actctgaggg 60
aggagctggg gacggcgacc ctaggagagt tctttggggg gactttcaag atggactcta 120
ctctaacagc aagtgaaatc cggcagcgat ttatagattt cttcaagagg aacgagcata 180
cgtatgttca ctcgtctgcc accatcccat tggatgaccc cactttgtct tttgccaatg 240
caggcatgaa ccagtttaaa cccatttttc tgaacacaat tgacccatct caccocatgg 300
caaagctgag cagagctgcc aatacccaga agtgcattcc ggctgggggc aaacaaaatg 360
acctggacga tgtgggcaag gatgtctatc atcacacctt cttcgagatg ctgggctctt 420
ggtcttttgg agattacttt aaggaattgg catgtaagat ggctctggaa ctctcacacc 480
aagagtttgg cattcccat tgaagacttt atgttactta ctttggcggg gatgaagcag 540
ctggcttaga agcagatctg gaatgcaaac agatctggca aaatttgggg ctggatgaca 600
ccaaaatcct cccaggcaac atgaaggata acttctggga gatgggtgac acggggcccct 660
gtggctcctg cagttagatc cactacgacc ggattgggtg tcgggacgcc gcacatcttg 720
tcaaccagga cgacccta atgtctggaga tctggaacct tgtgttcac cagtataaca 780
gggaagctga tggcattctg aaacctcttc ccaagaaaag cattgacaca gggatggggc 840
tggaacgact ggtatctgtg ctgcagaata agatgtccaa ctatgacact gacctttttg 900
tcccttactt tgaagccatt cagaaggggc cagggtgccc accatacact gggaaagtgt 960
gtgctgagga tgccgatggg attgacatgg cctaccgggt gctggctgac catgctcgga 1020
ccatcactgt ggcactgggt gatgggtggc ggcctgacaa cacagggcgt ggatatgtgt 1080
tgagacggat tctccgcoga gctgtccgat acgcccataa aaagctcaat gccagcaggg 1140
gcttctttgc tacgttagtg gatgttgtog tccagtcctt gggagatgca tttcctgagc 1200
tgaagaagga cccagacatg gtgaaggaca tcattaatga agaagaggtg cagttttctc 1260
agactctcag cagagggcgt cgcacacctt gcttggctcc tctatgacac ctatgggttt ccagtggatc 1320
agaccattcc cggagacact gcttggctcc tctatgacac ctatgggttt ccagtggatc 1380
tgactggact gattgctgaa gagaaggggc tgggtggtaga catggatggc tttgaagagg 1440
agaggaaact ggcccagctg aaatcacagg gcaagggagc tgggtggggaa gacctatta 1500
tgctggacat ttacgctatc gaagagctcc gggcacgggg tctggaggtc acagatgatt 1560
ccccaaagta caattaccat ttggactcca gtggttagct tgtatttgag aacacagtgg 1620
ctacggtgat ggctctgccc agggagaaga tgttcgtgga agaggtgtcc acagggcagg 1680
agtgtggagt ggtgctggac aagacctgtt tctatgctga gcaaggaggc cagatctatg 1740
acgaaggcta cctggtgaag gtggatgaca gcagtgaaga taaaacagag tttacagtga 1800
agaatgctca ggtccgagga gggatgtgtc tacacattgg aaccatctac ggtgacctga 1860
aagtggggga tcaggtctgg ctgtttattg atgagccccg acgaagacc atcatgagca 1920
accacacagc tacgcacatt ctgaacttcg ccctgcgctc agtgcttggg gaagctgacc 1980
agaaaggctc attggttgct cctgaccgcc tcagatttga ctttactgcc aaggagacca 2040
tgtccacca acagatcaag aaggctgaag agattgctaa tgagatgatt gaggcagcca 2100
aggccgtcta taccaggat tgccccctgg cagcagcgaa agccatccag ggcctacggg 2160
ctgtgtttga tgagacctat cctgaccctg tgcgagtcgt ctccattggg gtcccgggtg 2220
ccgagttgct ggatgacccc tctgggctct ctggctccct gacttctgtt gagttctgtg 2280
ggggaacgca cctgcggaac tccagtcatg caggagcttt tgtgatcgtg acggaagaag 2340
ccattgccaa gggatccggg aggatgtgtg ctgtcacagg tgccgaggcc cagaaggccc 2400
tcaggaaagc agagagcttg aagaaatgtc tctctgtcat ggaagccaaa gtgaaggctc 2460
agactgtctc aaacaaggat gtgcagagg agatcgctga ccttgagag gccctggcca 2520
ctgcagtcac ccccagtg cgagaaggat aattgcggga gactctcaaa tccctaaaga 2580
aggtcatgga tgacttggac cgagccagca aagccgatgt ccagaaacga gtgttagaga 2640

agacgaagca gttcatcgac agcaacccca accagcctct tgtcatcctg gagatggaga 2700
gccccgcctc agccaaggcc ctgaatgaag ccttgaagct cttcaagatg cactccccctc 2760

```



```

agactttctgc catgctcttc acggtggaca atgaggctgg caagatcacg tgccctgtgtc 2820
aagtccccca gaatgcagcc aatcggggct taaaagccag cgagtgggtg cagcaggtgt 2880
caggcttgat ggacggtaaa ggtggtggca aggatgtgtc tgcacaggcc acaggcaaga 2940
acgttggctg cctgcaggag gcgctgcagc tggccacttc cttcgcccag ctgcgcctcg 3000
gggatgtaaa gaactgagtg ggggaaggagg aggtccccc tggatccatc cgtccagcca 3060
agagctcttc atctgctaca agaacatttg aatcttggga cctttaaaga gcccctccta 3120
acccagcagt aactggaaca cacttgggag cagtccctatg tctcagtgcc ccttaaattt 3180
ctgccctgag ccctccacgt cagtgccatc ggtctagaac cactaacccc gcattgctgt 3240
tgatcgtcac gctcgcatct atagataacg gctctccaga cctgagcttt ccgctgcagc 3300
aagtaggaat cgtttttgct gcagagaata aaaggaccac gtgc 3344

```

<210> 78
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001605

<400> 78
 gccaaagagct cttcatctgc tacaagaaca tttgaatctt gggaccttta aagagcccct 60

<210> 79
 <211> 417
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001645

```

<400> 79
acctcccaac caagccctcc agcaaggatt caggagtgcc cctcgggcct cgccatgagg 60
ctcttctgtg cgtcccggt cctggtggtg gttctgtcga tcgtcttga aggccagcc 120
ccagcccagg ggacccaga cgtctccagt gccttgata agctgaagga gtttggaac 180
acactggagg acaaggctcg ggaactcatc agccgcatca aacagagtga actttctgcc 240
aagatgcggg agtggttttc agagacattt cagaaagtga aggagaaact caagattgac 300
tcatgaggac ctgaagggtg acatccagga ggggcctctg aaatttccca caccagcg 360
cctgtgctga ggactccgc catgtggccc cagggtgccac caataaaaat cctaccg 417

```

<210> 80
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001645

<400> 80
 aaacagagtg aactttctgc caagatgcgg gagtggtttt cagagacatt tcagaaagtg 60

<210> 81
 <211> 1389
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001809

<400> 81
 cgcggacttc tgccaagcac cggctcatgt gaggctcgcg gcacagcgtt ctctgggctc 60

```

cccagaagcc agccttttgc tcccggaccc ggcagcccca gcaggagccg tgggaccggg 120
cgccagcacc ctctgcggcg tgtcatgggc ccgcgcgcgc ggagccgaaa gcccagggcc 180
ccgaggaggc gcagcccagc cccgaccccg acccccgggc cctcccggcg gggccctcc 240
ttaggcgctt cctcccatca acacagtcgg cggagacaag gttggctaaa ggagatccga 300
aagcttcaga agagcacaca cctcttgata aggaagctgc ccttcagccg cctggcaaga 360
gaaatatgtg ttaaattcac tcgtggtgtg gacttcaatt ggcaagccca ggccctattg 420
gccctacaag aggcagcaga agcattttcta gttcatctct ttgaggacgc ctatctcctc 480
accttacatg caggccgagt tactctcttc ccaaaggatg tgcaactggc ccggaggatc 540
cggggccttg aggagggact cggctgagct cctgcaccca gtgtttctgt cagtctttcc 600
tgctcagcca ggggggatga taccggggac tctccagagc catgactaga tccaatggat 660
tctgcgatgc tgtctggact ttgctgtctc tgaacagtat gtgtgtgttg ctttaaata 720
ttttcttttt tttgagaagg agaagactgc atgactttcc tctgtaacag aggtaatata 780
tgagacaatc aacaccgttc caaaggcctg aaaataattt tcagataaag agactccaag 840
gttgacttta gtttgtgagt tactcatgtg actatttgag gattttgaaa acatcagatt 900
tgctgtggta tgggagaaaa ggttatgtac ttattatttt agctctttct gtaatattta 960
cattttttac catatgtaca tttgtacttt tatttttacac ataagggaaa aaataagacc 1020
actttgagca gttgcctgga aggctgggca tttccatcat atagacctct gcccttcaga 1080
gtagcctcac cattagtggc agcatcatgt aactgagtgg actgtgcttg tcaacggatg 1140
tgtagctttt cagaaactta attggggatg aatagaaaac ctgtaagctt tgatgttctg 1200
gttactttcta gtaaattcct gtcaaaatca attcagaaat tctaacttgg agaatttaac 1260
attttactct tgtaaatacat agaagatgta tcataacagt tcagaatttt aaagtacatt 1320
ttcgatgctt ttatgggtat tttttagtct tctttgtaga gagataataa aaatcaaaat 1380
atthaatga 1389

```

<210> 82

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001809

<400> 82

```

ggggatgaat agaaaacctg taagctttga tgttctgggt acttctagta aattcctgtc 60

```

<210> 83

<211> 2205

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001909

<400> 83

```

gcgcacgccg gccgcgcca cgtgaccggg ccgggtgcaa acacgggggt cagctgatcc 60
ggcccaactg cggcgtcatc ccggctataa gcgcacgggc tgggcgacc tctccgacc 120
ggccgcgcgc gccatgcagc cctccagcct tctgccgctc gccctctgcc tgcctggctgc 180
accgcctccc gcgctcgtca ggatcccgtc gcacaagtcc acgtccatcc gccggaccat 240
gtcggagggtt gggggctctg tggaggacct gattgccaaa ggccccgtct caaagtactc 300
ccaggcggtg ccagccgtga ccgagggggc cattcccagc gtgctcaaga actacatgga 360
cgcccagtac tacggggaga ttggcatcgg gacgcccccc cagtgettca cagtctctt 420
cgacacgggc tctccaacc tgtgggtccc ctccatccac tgcaaactgc tggacatgc 480
ttgctggatc caccacaagt acaacagcga caagtccagc acctacgtga agaattggat 540
ctcgtttgac atccactatg gctcgggcag cctctccggg tacctgagcc aggacactgt 600
gtcggtgccc tgccagtcag cgtcgtcagc ctctgccctg ggcgggtgtc aagtggagag 660
gcaggtcttt ggggaggcca ccaagcagcc aggcataacc ttcatcgcag ccaagttcga 720
tggtcatcctg ggcattggcct acccccgcct ctccgtcaac aacgtgctgc cgtcttccga 780
caacctgatg cagcagaagc tgggtggacca gaacatcttc tccttctacc tgagcagggg 840
cccagatgcg cagcctgggg gtgagctgat gctgggtggc acagactcca agtattacaa 900

```

```

gggtttctctg tectacctga atgtcacccg caaggcctac tggcaggtcc acctggacca 960
ggtggagggtg gccagcgggc tgacctgtg caaggagggc tgtgaggcca ttgtggacac 1020
aggcacttcc ctcatggtgg gcccggtgga tgagggtgcg gagctgcaga aggccatcgg 1080
ggccgtgcgg ctgattcagg gcgagtacat gatccccctgt gagaaggtgt ccacctgcc 1140
cgcgatcaca ctgaagctgg gaggcaaagg ctacaagctg tccccagagg actacacgct 1200
caaggtgtcg caggccggga agacctctg cctgagcggc ttcattgggca tggacatccc 1260
gccacccagc gggccactct ggatcctggg cgacgtcttc atcgcccgct actacactgt 1320
gtttgacctg gacaacaaca ggggtgggctt cgccgaggtt gcccgctctt agttcccaag 1380
gcgtccgcgc gccagcacag aaacagagga ggtcccaga gcaggaggcc cctggcccag 1440
cggccctccc cacacacacc cacacactcg ccgcccact gtccctgggcg cctgggaagc 1500
cggcggccca agcccactt gctgttttgt tctgtgggtt tcccctccct gggttcagaa 1560
atgctgcctg cctgtctgtc tctccatctg tttgggtggg gtagagctga tccagagcac 1620
agatctgttt cgtgcattgg aagacccac ccaagcttgg cagccgagct cgtgtatcct 1680
ggggctccct tcatctccag ggagtcacct ccccgccct accagcgccc gctgggctga 1740

```

```

gcccctaccc cacaccaggc cgtcctcccg ggccctccct tggaaacctg cctgcctga 1800
gggcccctct gccagcttg ggccagctg ggctctgcca ccctacctgt tcagtgtccc 1860
gggcccgttg aggatgaggc cgctagaggc ctgaggatga gctggaagga gtgagagggg 1920
acaaaaccca ccttgtttga gcctgcaggg tgggtgctggg actgagccag tcccaggggc 1980
atgtattggc ctggaggtgg ggttgggatt gggggctggt gccagccttc ctctgcagct 2040
gacctctgtt gtccctccct tggggcggtg agagccccag ctgacatgga aatacagttg 2100
ttggcctccg gcctccctc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2160
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 2205

```

<210> 84
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001909

```

<400> 84
tctgtttggt gggggtagag ctgatccaga gcacagatct gtttcgtgca ttggaagacc 60

```

<210> 85
<211> 817
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002038

```

<400> 85
gaaccgttta ctcgctgctg tgcccatcta tcagcaggct ccgggctgaa gattgcttct 60
cttctctcct ccaaggctcta gtgacggagc ccgcgcgcgg cgccaccatg cggcagaagg 120
cggtatcgct tttcttgtgc tacctgctgc tcttcaactg cagtgggggtg gaggcaggta 180
agaaaaagtg ctcgagagac tcggacagcg gctccgggtt ctggaaggcc ctgaccttca 240
tgcccgctcg aggaggactc gcagtcgccc ggctgcccgc gctgggcttc accggcgccg 300
gcatecgggc caactcggtg gctgectcgc tgatgagctg gtctgcgata ctgaatgggg 360
gcggcggtgc cgccgggggg ctagtggcca cgctgcagag cctcggggct ggtggcagca 420
gcgtcgatcat aggtaatatt ggtgccctga tgggctacgc caccacaaag tatctcgata 480
gtgaggagga tgaggagtag ccagcagctc ccagaacctc ttcttccttc ttggcctaac 540
tcttccagtt aggatctaga actttgcctt tttttttttt tttttttttt tttgagatgg 600
gttctcacta tattgtccag gctagagtgc agtggctatt cacagatgcg aacatagtag 660
actgcagcct ccaactcta gectcaagt atcctcctgt ctcaacctcc caagtaggat 720
tacaagcatg cgccgacgat gcccagaatc cagaactttg tctatcactc tccccaaaca 780
cctagatgtg aaaacagaat aaacttcacc cagaaaa 817

```

<210> 86
<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_002038

<400> 86
agctcccaga acctcttctt ccttcttggc ctaactcttc cagttaggat ctagaacttt 60

<210> 87
<211> 1283
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002046

<400> 87
ctctctgctc ctctctgttcg acagtcagcc gcatcttctt ttgcgtcgcc agccgagcca 60
catcgctcag acaccatggg gaaggtgaag gtcggagtca acggatttgg tcgtattggg 120
cgctgtgtca ccagggctgc ttttaactct ggtaaagtgg atattgttgc catcaatgac 180
cccttcattg acctcaacta catggtttac atgttccaat atgattccac ccatggcaaa 240
ttccatggca ccgtcaaggc tgagaacggg aagcttgtca tcaatggaaa tcccatcacc 300
atcttccagg agcgagatcc ctccaaaatc aagtggggcg atgctggcgc tgagtacgtc 360
gtggagtcca ctggcgtctt caccaccatg gagaaggctg gggctcattt gcagggggga 420
gccaaaaggg tcacatctct tgccccctct gctgatgccc ccatgttcgt catgggtgtg 480
aaccatgaga agtatgacaa cagcctcaag atcatcagca atgcctcctg caccaccaac 540
tgcttagcac ccctggccaa ggtcatccat gacaactttg gtatcgtgga aggactcatg 600
accacagtcc atgccatcac tgccacccag aagactgtgg atggccctc cgggaaactg 660
tggcgtgatg gccgcggggc tctccagaac atcatccctg cctctactgg cgctgccaaag 720
gctgtgggca aggtcatccc tgagctgaac ggggaagctc ctggcatggc cttccgtgtc 780
cccactgcc aagtgtcagt ggtggacctg acctgccgtc tagaaaaacc tgccaaatat 840
gatgacatca agaagggtgg gaagcaggcg tcggaggggc cctcaagggt catcctgggc 900
tacactgagc accagggtgg ctctcttgac ttcaacagcg acacccactc ctccacctt 960
gacgtgggg ctggcattgc cctcaacgac cactttgtca agctcatttc ctggtatgac 1020
aacgaatttg gctacagcaa cagggtgggtg gacctcatgg ccacatggc ctccaaggag 1080
taagaccctt ggaccaccag ccccagcaag agcacaagag gaagagagag accctcactg 1140
ctggggagtc cctgccacac tcagtcccc accacactga atctccctc ctacacagttg 1200
ccatgtagac cccttgaaga ggggaggggc ctaggggagcc gcaccttgtc atgtaccatc 1260
aataaagtac cctgtgctca acc 1283

<210> 88
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002046

<400> 88
ctcaacgacc actttgtcaa gctcatttcc tggtatgaca acgaatttgg ctacagcaac 60

<210> 89
<211> 1610
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002061

<400> 89

```

ggcaccgaggc tgcgggccgca gtagccgggag ccggagccgc agccaccggt gccttccttt 60
cccgccgccc cccagccgcc gtccggccctc cctcggggccc gagcgagac caggctccag 120
ccgcgcggcg ccggcagcct cgcgctccct ctogggtctc tctcgggctt cgggcaccgc 180
gtcctgtggg cggccgcctg cctgcccgcg cgcccgcagc cccttgcttg cgggcccctg 240
ggcggcccggt gccatgggca ccgacagccg cgcggccaag gcgctcctgg cgcgggcccg 300
caccctgcac ctgcagacgg ggaacctgct gaactggggc cgcctgcgga agaagtgcc 360
gtccacgcac agcgaggagc ttcatgattg tatccaaaaa accttgaatg aatggagttc 420
ccaaatcaac ccagatttgg tcagggagtt tccagatgtc ttggaatgca ctgtatctca 480
tgcagtagaa aagataaatc ctgatgaaag agaagaaatg aaagtttctg caaaactgtt 540
cattgtagaa tcaaactctt catcatcaac tagaagtgcg gttgacatgg cctgttcagt 600

```

```

ccttgaggatt gcacagctgg attctgtgat cattgcttca cctcctattg aagatggagt 660
taatctttcc ttggagcatt tacagcctta ctgggaggaa ttagaaaact tagttcagag 720
caaaaagatt gttgccatag gtacctctga tctagacaaa acacagttgg aacagctgta 780
tcagtgggca caggtaaaac caaatagtaa ccaagttaat cttgcctcct gctgtgtgat 840
gccaccagat ttgactgcat ttgctaaaca atttgacata cagctgttga ctcacaatga 900
tccaaaagaa ctgctttctg aagcaagttt ccaagaagct cttcaggaaa gcattcctga 960
cattcaagcg cagcagtggtg tgccgctgtg gctactgcgg tattcgggtc ttgtgaaaag 1020
tagaggaatt atcaaatcaa aaggctacat tttaacaagc aaaagaaggg gttcttaact 1080
gacttaggag cataacttac ctgtaatttc cttcaatatg agagaaaatt gagatgtgta 1140
aaatctagtt actgcctgta aatgggtgtc ttgaggcaga tattctttctg tcataattga 1200
cagtatgttg tctgtcaagt tttaaatact tatcttgcct ccataatcaat ccattctcat 1260
gaacctctgt attgctttcc ttaaactatt gttttctaata tgaaattgtc tataaagaaa 1320
atacttgcaa tatatttttc ctttattttt atgactaata taaatcaaga aaatttgttg 1380
ttagatatat tttggcctag gtatcagggt aatgtatata catatttttt atttccaaaa 1440
aaaattcatt aattgcttct taactcttat tataaccaag caatttaatt acaattgtta 1500
aaactgaaat actggaagaa gatatttttc ctgtcattga tgagatatat cagagtaact 1560
ggagtagctg ggatttacta gtagtgtaaa taaaattcac tcttcaatac 1610

```

<210> 90

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002061

<400> 90

```

ctgacttagg agcataactt acctgtaatt tccttcaata tgagagaaaa ttgagatgtg 60

```

<210> 91

<211> 873

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002106

<400> 91

```

cgcagtttga atcgcggtgc gacgaaggag taggtggtgg gatctcaccg tgggtccgat 60
tagccttttc tctgccttgc ttgcttgagc ttcagcggaa ttcgaaatgg ctggcggtta 120
ggctggaaaag gactccggaa aggccaagac aaaggcgggt tcccgcctgc agagagccgg 180
cttgagtttc ccagtggggc gtattcatcg acacctaaaa tctaggacga ccagtcagtg 240
acgtgtggggc gcgactgccg ctgtgtacag cgcagccatc ctggagtaac tcaccgcaga 300
ggtacttgaa ctggcaggaa atgcatcaaa agacttaag gttaaagcgta ttaccctctg 360
tcacttgcaa cttgctattc gtggagatga agaattggat tctctcatca aggctacaat 420
tgctggtggg ggtgtcattc cacacatcca caaatctctg attgggaaga aaggacaaca 480
gaagactgtc taaaggatgc ctggattcct tgttatctca ggactctaaa tactctaaca 540
gctgtccagt gttggtgatt ccagtggact gtatctctgt gaaaaacaca attttgctt 600
tttgtaattc tatttgagca agttggaagt ttaattagct ttccaaccaa ccaaatttct 660

```

```
gcattcgagt ctttaaccata ttttaagtgtt actgtggcctt caaagaagct attgattctg 720
aagtagtggg ttttgattga gttgactgtt ttttaaaaaac tgtttggatt ttaattgtga 780
tgcagaagtt atagtaacaa acatttggtt ttgtacagac attattttcca ctctggtgga 840
taagttcaat aaaggtcata tcccaaacta aaa 873
```

```
<210> 92
<211> 60
<212> DNA
<213> Homo sapiens
```

```
<300>
<308> NM_002106
```

```
<400> 92
cgagtcttaa ccatatttaa gtgttactgt ggcttcaaag aagctattga ttctgaagta 60
```

```
<210> 93
<211> 4204
<212> DNA
<213> Homo sapiens
```

```
<300>
<308> NM_002205
```

```
<400> 93
caggacaggg aagagcgggc gctatgggga gccggacgcc agagtccctt ctccacgccg 60
tgcagctgcg ctggggccccc cggcgccgac ccccgctcgt gccgctgctg ttgctgctcg 120
tgccgccgcc acccagggtc gggggccttca acttagacgc ggaggcccca gcagtactct 180
cggggccccc gggctccttc ttoggattct cagtggagtt ttaccggccg ggaacagacg 240
gggtcagtgt gctggtggga gcacccaagg ctaataccag ccagccagga gtgctgcagg 300
gtggtgctgt ctacctctgt ccttgggggtg ccagcccccac acagtgcacc cccattgaat 360
ttgacagcaa aggtctctcg ctctggaggt cctcactgtc cagctcagag ggagaggagc 420
ctgtggagta caagtccttg cagtgggttcg gggcaacagt tcgagcccat ggctcctcca 480
tcttggcatg cgctccactg tacagctggc gcacagagaa ggagccactg agcgacccccg 540
tgggcacctg ctacctctcc acagataact tcaccogaat tctggagtat gcacctgcc 600
gctcagattt cagctgggca gcaggacagg gttactgccca aggaggcttc agtgccaggt 660
tcaccaagac tggccgtgtg gttttaggtg gaccaggaag ctatttctgg caaggccaga 720
tcctgtctgc cactcaggag cagattgcag aatcttatta ccccgagtac ctgatcaacc 780
tggttcaggg gcagctgcag actcgccagg ccagttccat ctatgatgac agctacctag 840
gatactctgt ggctgttggg gaattcagtg gtgatgacac agaagacttt gttgctgggtg 900
tgcccaaagg gaacctcact tacggctatg tcaccatcct taatggctca gacattogat 960
ccctctacaa cttctcaggg gaacagatgg cctcctactt tggctatgca gtggccgcc 1020
cagacgtcaa tggggacggg ctggatgact tgcctggtggg ggcacccctg ctcatggatc 1080
ggacccctga cgggcgccct caggaggtgg gcagggtcta cgtctacctg cagcaccacg 1140
ccggcataga gccacgccc acccttacct tcactggcca tgatgagttt ggccgatttg 1200
gcagctcctt gacccccctg ggggacctgg accagratgg ctacaatgat gtggccatcg 1260
gggctccctt tgggtggggag acccagcagg gagttagtgt tgtatttctt gggggccacg 1320
gagggctggg ctctaagcct tcccaggttc tgcagccct gtgggcagcc agccacaccc 1380
cagacttctt tggctctgcc cttcgaggag gccgagacct ggatggcaat ggatatcctg 1440
atctgattgt ggggtccttt ggtgtggaca aggtgtggtg atacaggggc cgccccatcg 1500
tgtccgctag tgctccctc accatcttcc ccgccatgtt caaccacagag gagcggagct 1560
gcagcttaga ggggaaccct gtggcctgca tcaaccttag ctctgcctc aatgcttctg 1620
gaaaacacgt tgctgactcc attggtttca cagtggaaact tcagctggac tggcagaagc 1680
agaagggagg ggtacggcgg gcactgttcc tggcctccag gcaggcaacc ctgaccaga 1740
ccctgctcat ccagaatggg gctcgagagg attgcagaga gatgaagatc tacctcagga 1800
acgagtcaga atttcgagac aaactctcgc cgattcacat cgctctcaac ttctccttgg 1860
acccccaaag cccagtggac agccacggcc tcaggccagc cctacattat cagagcaaga 1920
gccggataga ggacaaggct cagatcttgc tggactgtgg agaagacaac atctgtgtgc 1980
ctgacctgca gctggaagtg tttggggagc agaaccatgt gtacctgggt gacaagaatg 2040
ccctgaacct cactttccat gcccagaatg tgggtgaggg tggcgctat gaggctgagc 2100
ttcgggtcac cgccccctca gaggtctgagt actcaggact cgtcagacac ccagggaact 2160
```

```

tctccagcct gagctgtgac tactttgccc tgaaccagag ccgcctgctg gtgtgtgacc 2220
tgggcaaccc catgaaggca ggagccagtc tgtgggggtgg ccttcgggtt acagtccctc 2280
atctccggga cactaagaaa accatccagt ttgacttcca gatcctcagc aagaatctca 2340
acaactcgca aagcgacgtg gtttcccttc ggctctccgt ggaggctcag gccaggtca 2400
ccctgaacgg tgtctccaag cctgaggcag tgctattccc agtaagcgac tggcatcccc 2460
gagaccagcc tcagaaggag gaggaacctg gacctgctgt ccaccatgtc tatgagctca 2520
tcaaccaagg cccagctcc attagccagg gtgtgctgga actcagctgt cccagggctc 2580
tggaagggtca gcagctccta tatgtgacca gagttacggg actcaactgc accaccaatc 2640
acccattaa cccaaagggc ctggagttgg atcccagggg ttccctgcac caccagcaaa 2700
aacgggaagc tccaagccgc agctctgctt cctcgggacc tcagatcctg aaatgcccg 2760
aggctgagtg tttcaggctg cgctgtgagc tcgggcccct gcaccaacaa gagagccaaa 2820
gtctgcagtt gcatttccga gtctgggcca agactttctt gcagcgggag caccagccat 2880
ttagcctgca gtgtgaggct gtgtacaaag cctgaagat gccctaccga atcctgctc 2940
ggcagctgcc ccaaaaagag cgtcaggtgg ccacagctgt gcaatggacc aaggcagaag 3000
gcagctatgg cgtcccactg tggatcatca tcttagccat cctgtttggc ctctgctcc 3060
taggtctact catctacatc ctctacaagc ttggattctt caaacgctcc ctcccatatg 3120
gcaccgccat ggaaaaagct cagctcaagc ctccagccac ctctgatgcc tgagtcctcc 3180
caatttcaga ctcccattcc tgaagaacca gtccccccac cctcattcta ctgaaaagga 3240
ggggtctggg tacttcttga aggtgctgac ggccaggagg aagctcctct cccagccca 3300
gagacatact tgaagggccca gagccagggg ggtgaggagc tggggatccc tccccccat 3360
gcactgtgaa ggacccttgt ttacacatac cctcttcatg gatgggggaa ctcagatcca 3420
gggacagagg cccagcctcc ctgaagcctt tgcattttgg agagtttcct gaaacaactg 3480
gaaagataac taggaaatcc attcacagtt ctttgggcca gacatgccac aaggacttcc 3540
tgtccagctc caacctgcaa agatctgtcc tcagccttgc cagagatcca aaagaagccc 3600
ccagtaagaa cctggaactt ggggagttaa gacctggcag ctctggacag cccaccctg 3660
gtggggcaac aaagaacact aactatgcat ggtgccccag gaccagctca ggacagatgc 3720
cacaaggata gatgctggcc cagggccaga gccagctcc aaggggaatc agaactcaaa 3780
tggggccaga tccagcctgg ggtctggagt tgatctggaa ccagactca gacattggca 3840

```

```

ccaatccagg cagatccagg actatatttg ggcctgctcc agacctgac ctggaggccc 3900
agttcacccct gatttaggag aagccaggaa tttcccagga cctgaagggg ccatgatggc 3960
aacagatctg gaacctcagc ctggccagac acaggccctc cctgttcccc agagaaagg 4020
gagccactg tcttgggctc gcagaatttg ggttctgcct gccagctgca ctgatgctgc 4080
ccctcatctc tctgcccac ccttccctca ccttggcacc agacaccag gacttattta 4140
aactctgttg caagtgcaat aaatctgacc cagtgcctcc actgaccaga actagaaaaa 4200
aaaa 4204

```

<210> 94
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002205

<400> 94
 ttggcaccag acaccagga cttattttaa ctctgttgca agtgcaataa atctgaccca 60

<210> 95
 <211> 1976
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002266

<400> 95
 gccacacggg ctttgagctg agtcgagggt gaccctttga acgcagtcgc cctacagccg 60
 ctgattcccc ccgcacgcgc tcccgtggaa gccaggccc gcttcgcagc tttctccctt 120
 tgtctcataa ccatgtccac caacgagaat gctaatacac cagctgcccg tcttcacaga 180
 ttcaagaaca agggaaaaga cagtacagaa atgaggcgtc gcagaataga ggtcaatgtg 240

```

gagctgagga aagctaagaa ggatgaccag atgctgaaga ggagaaatgt aagctcattt 300
cctgatgatg ctactttctcc gctgcaggaa aaccgcaaca accagggcac tgtaaattgg 360
tctgttgatg acattgtcaa aggcataaat agcagcaatg tggaaaatca gctccaagct 420
actcaagctg ccaggaaact actttccaga gaaaaacagc ccccataga caacataatc 480
cgggctgggtt tgattccgaa atttgtgtcc ttcttgggca gaactgattg tagtcccatt 540
cagtttgaat ctgcttgggc actcactaac attgcttctg ggacatcaga acaaaaccaag 600
gctgtggtag atggaggtgc catcccagca ttcattttctc tgttggcatc tccccatgct 660
cacatcagtg aacaagctgt ctgggctcta ggaaacattg caggtgatgg ctcagtgctc 720
cgagacttgg ttattaagta cgggtgcagtt gaccactgt tggctctcct tgcagttcct 780
gatatgtcat ctttagcatg tggctactta cgtaatctta cctggacact ttctaattct 840
tgccgcaaca agaatcctgc acccccgata gatgctgttg agcagattct tcctacctta 900
gttcggctcc tgcacatga tgatccagaa gtgttagcag atacctgctg ggctatttcc 960
taccttactg atggtccaaa tgaacgaatt ggcattggtg tgaaaacagg agttgtgccc 1020
caacttgtga agcttctagg agcttctgaa ttgccaattg tgactcctgc cctaagagcc 1080
atagggaata ttgtcactgg tacagatgaa cagactcagg ttgtgattga tgcaggagca 1140
ctcgccgtct ttcccagcct gctcaccaac cccaaaacta acattcagaa ggaagctacg 1200
tggacaatgt caaacatcac agccggccgc caggaccaga tacagcaagt tgtgaatcat 1260
ggattagtc cttccttctg cagtgttctc tctaaggcag attttaagac acaaaaggaa 1320
gctgtgtggg ccgtgaccaa ctataccagt ggtggaacag ttgaacagat tgtgtacctt 1380
gttcaactgt gcataataga accgttgatg aacctcttaa ctgcaaaaga taccaagatt 1440
attctggtta tcttggtatg catttcaaat atctttcagg ctgctgagaa actaggtgaa 1500
actgagaaac ttagtataat gattgaagaa tgtggaggct tagacaaaat tgaagctcta 1560
caaaaccatg aaaatgagtc tgtgtataag gcttcgttaa gcttaattga gaagtatttc 1620
tctgtagagg aagaggaaga tcaaaacggt gtaccagaaa ctacctctga aggctacact 1680
ttccaagttc aggatggggc tctggggacc tttaactttt agatcatgta gctgagacat 1740
aaatttgttg tgtactacgt ttggtatttt gtcttattgt ttctctacta agaactcttt 1800
cttaaatgtg gtttgttact gtagcacttt ttacactgaa actatacttg aacagttcca 1860
actgtacata catactgtat gaagcttgct ctctgactag gtttctaatt tctatgtgga 1920
atttcctatc ttgcagcatc ctgtaaataa acattcaagt ccacccttaa aaaaaa 1976

```

<210> 96
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002266

<400> 96
 tgagtctgtg tataaggctt cgtaaagctt aattgagaag tatttctctg tagaggaaga 60

<210> 97
 <211> 1145
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002346

<400> 97
 gctccggcca gccgcgggtcc agagcgcgcg aggttcgggg agctccgcca ggctgctggt 60
 acctgctgcc gcccggcgag caggacaggc tgctttgggt tgtgaacctcc aggcaggacg 120
 gccatcctct ccagaatgaa gatcttcttg ccagtgtctc tggctgccct tctgggtgtg 180
 gagcgagcca gctcgtgat gtgcttctcc tgcctgaacc agaagagcaa tctgtactgc 240
 ctgaagccga ccatctgctc cgaccaggac aactactgcy tgactgtgtc tgctagtgcc 300
 ggcatgtgga atctcgtgac atttggccac agcctgagca agacctgttc cccggcctgc 360
 cccatcccag aaggcgtcaa tgttgggtgtg gcttccatgg gcatcagctg ctgccagagc 420
 tttctgtgca atttcagtgc ggccgatggc gggctgcggg caagcgtcac cctgctgggt 480
 gccgggctgc tgcagagcct gctgccggcc ctgctgcggg ttggcccttg accgcccaga 540
 cctgtcccc cgatccccc gctcaggaag gaaagcccag cctttcttg atccacagt 600
 gtatgggagc cctgactcc tcactgtcct gatctgtgcc cttggtccca ggtcaggccc 660


```

acccccctgca cctccacctg ccccagcccc tgcctctgcc caagtggggc agctgccctc 720
acttctggggg tggatgatgt gaccttccct gggggactgc ggaaggacg agggttccct 780
ggagtctttac ggtccaacat cagaccaagt cccatggaca tgctgacagg gtccccaggg 840
agaccgtgtc agtagggatg tgtgcctggc tgtgtacgtg ggtgtgcagt gcacgtgaga 900
gcacgtggcg gcttctgggg gccatgtttg gggaggaggg tgtgccagca gcctggagag 960
cctcagtccc tgtagcccc tgcctggca cagctgcagt cacttcaagg gcagcctttg 1020
ggggttgggg tttctgccac ttccgggtct aggcctgcc caaatccagc cagtccctgcc 1080
ccagcccacc cccacattgg agccctcctg ctgctttggt gcctcaaata aatacagatg 1140
tcccc 1145

```

```

<210> 98
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002346

```

```

<400> 98
ggttccctgg agtcttacgg tccaacatca gaccaagtcc catggacatg ctgacagggt 60

```

```

<210> 99
<211> 1390
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002358

```

```

<400> 99
gggaagtgct gttggagccg ctgtggttgc tgtccgcgga gtggaagcgc gtgcttttgt 60
ttgtgtccct ggccatggcg ctgcagctct cccgggagca gggaatcacc ctgcgcggga 120
gcgccgaaat cgtggccgag ttcttctcat tcggcatcaa cagcatttta tatcagcgtg 180
gcataatatc atctgaaacc tttactcgag tgcagaaata cggactcacc ttgcttgtaa 240
ctactgatct tgagctcata aaatacctaa ataatgtggt ggaacaactg aaagattggt 300
tatacaagtg ttcagttcag aaactggttg tagttatctc aaatattgaa agtgggtgagg 360
tcctggaaag atggcagttt gatattgagt gtgacaagac tgcaaaagat gacagtgcac 420
ccagagaaaa gtctcagaaa gctatccagg atgaaatccg ttcagtgate agacagatca 480
cagctacggg gacatttctg ccactgttgg aagtttcttg ttcatttgat ctgctgattt 540
atacagacaa agatttggtt gtacctgaaa aatgggaaga gtcgggacca cagtttatta 600
ccaattctga ggaagtcgcg cttcgttcat ttactactac aatccacaaa gtaaatagca 660
tggtggccta caaaattcct gtcaatgact gaggatgaca tgaggaaaat aatgtaattg 720
taattttgaa atgtggtttt cctgaaatca ggtcatctat agttgatatg ttttatttca 780
ttgggttaatt ttacatgga gaaaaccaa atgatactta ctgaactgtg tgtaattggt 840
cctttatttt tttggtacct atttgactta ccatggagtt aacatcatga atttattgca 900
cattgttcaa aaggaaccag gaggtttttt tgtcaacatt gtgatgtata ttctttgaa 960
gatagtaact gtagatggaa aaacttgtgc tataaagcta gatgctttcc taaatcagat 1020
gttttgggtca agtagtttga ctcagtatag gtaggagat atttaagtat aaaatacaac 1080
aaaggaagtc taaatattca gaatctttgt taaggtcctg aaagtaactc ataactata 1140
aacaatgaaa tattgctgta tagctccttt tgaccttcat ttcatgtata gttttcccta 1200
ttgaatcagt ttccaattat ttgactttta tttatgtaac ttgaacctat gaagcaatgg 1260
atatttgtac tgtttaatgt tctgtgatac agaactctta aaaatgtttt ttcatgtgtt 1320
ttataaaatc aagttttaag tgaaagtgag gaaataaagt taagtttgtt ttaaaaaaaaa 1380
aaaaaaaaa 1390

```

```

<210> 100
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>

```

<308> NM_002358

<400> 100

atgctttcct aaatcagatg ttttgggtcaa gtagtttgac tcagttatagg tagggagata 60

<210> 101

<211> 1821

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002422

<400> 101

```

acaaggaggc aggcaagaca gcaaggcata gagacaacat agagctaagt aaagccagtg 60
gaaatgaaga gtcttccaat cctactgttg ctgtgcgtgg cagtttgctc agcctatcca 120
ttggatggag ctgcaagggg tgaggacacc agcatgaacc ttgttcagaa atatctagaa 180
aactactacg acctcaaaaa agatgtgaaa cagtttggtta ggagaaaagga cagtggctct 240
gttgttaaaa aaatccgaga aatgcagaag ttccttggat tggagggtgac ggggaagctg 300
gactccgaca ctctggaggt gatgcgcaag ccaggtgtg gagttcctga tgttgggtcac 360
ttcagaacct ttcctggcat cccgaagtgg aggaaaaccc accttacata caggattgtg 420
aattatacac cagatttggc aaaagatgct gttgattctg ctgttgagaa agctctgaaa 480
gtctgggaag aggtgactcc actcacattc tccaggctgt atgaaggaga ggctgatata 540
atgatctctt ttgcagttag agaacatgga gacttttacc cttttgatgg acctggaaat 600
gttttggccc atgcctatgc ccctgggcca gggattaatg gagatgccc ctttgatgat 660
gatgaacaat ggacaaagga tacaacaggg accaatttat ttctcgttgc tgctcatgaa 720
attggccact ccctgggtct ctttactca gccaacactg aagctttgat gtaccactc 780
tatcactcac tcacagacct gactcggttc cgcctgtctc aagatgatat aaatggcatt 840
cagtcctctt atggacctcc ccctgactcc cctgagacct ccctgggtacc cacggaacct 900
gtccctccag aacctgggac gccagccaac tgtgatcctg ctttgtcctt tgatgctgtc 960
agcactctga ggggagaaat cctgatcttt aaagacaggc acttttggcg caaatccctc 1020
aggaagcttg aacctgaatt gcatttgatc tcttcatttt ggccatctct tccttcaggc 1080
gtggatgccg catatgaagt tactagcaag gacctcgttt tcatttttaa aggaaatcaa 1140
ttctgggcca tcagaggaaa tgaggtagca gctggatacc caagaggcat ccacacccta 1200
ggtttccctc caaccgtgag gaaaatcgat gcagccattt ctgataagga aaagaacaaa 1260
acatatttct ttgtagagga caaatactgg agatttgatg agaagagaaa ttccatggag 1320
ccaggctttc ccaagcaaat agctgaagac tttccaggga ttgactcaaa gattgatgct 1380
gtttttgaag aatttgggtt cttttatttc tttactggat cttcacagtt ggagtttgac 1440
ccaaatgcaa agaaagtga acacactttg aagagtaaca gctggcttaa ttgttgaaag 1500
agatatgtag aaggcacaat atgggcactt taaatgaagc taataattct tcacctaaag 1560
ctctgtgaat tgaaatgttc gttttctcct gcctgtgctg tgactcgagt cacactcaag 1620
ggaacttgag cgtgaatctg tatcttgccg gtcattttta tgttattaca gggcattcaa 1680

atgggctgct gcttagcttg caccttgtoa catagagtga tctttcccaa gagaagggga 1740
agcactcgtg tgcaacagac aagtgactgt atctgtgtag actatttgct tatttaataa 1800
agacgatttg tcagttgttt t 1821

```

<210> 102

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002422

<400> 102

tgtagaaggc acaatatggg cactttaaat gaagctaata attcttcacc taagtctctg 60

<210> 103

<211> 2787

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002462

<400> 103

```

agagcggagg cgcactcca gactgcgca gggaccgcct tggaccgcag ttgccggcca 60
ggaatcccag tgtcacggtg gacacgcctc cctcgcgccc ttgccggcca cctgctcacc 120
cagctcaggg gctttggaat tctgtggcca cactgcgagg agatcggttc tgggtcggag 180
gctacaggaa gactcccact ccctgaaatc tggagtgaag aacgccgcca tccagccacc 240
attccaagga ggtgcaggag aacagctctg tgataccatt taacttggtg acattacttt 300
tatttgaagg aacgtatatt agagcttact ttgcaaagaa ggaagatggt tgtttccgaa 360
gtggacatcg caaaagctga tccagctgct gcatcccacc ctctattact gaatggagat 420
gctactgtgg cccagaaaaa tccaggctcg gtggctgaga acaacctgtg cagccagtat 480
gaggagaagg tgcgcccctg catcgacctc attgactccc tgcgggctct aggtgtggag 540
caggacctgg ccctgccagc catcgccgtc atcggggacc agagctcggg caagagctcc 600
ctggttgagg cactgtcagg agttgccctt cccagaggca gcgggatcgt gaccagatgc 660
ccgctgggtg tgaaactgaa gaaacttggt aacgaagata agtggagagg caaggtcagt 720
taccaggact acgagattga gatttcggat gcttcagagg tagaaaagga aattaataaa 780
gcccagaatg ccacgcggg ggaaggaatg ggaatcagtc atgagctaata caccctggag 840
atcagctccc gagatgtccc ggatctgact ctaatagacc ttcctggcat aaccagagt 900
gctgtgggca atcagcctgc tgacattggg tataagatca agacactcat caagaagtac 960
atccagaggg aggagacaat cagcctgggt gtggtcccc gtaatgtgga catcgccacc 1020
acagaggtct tcagcatggc ccaggagggt gaccccgagg gagacaggac catcggaatc 1080
ttgacgaagc ctgactctgg ggacaaagga actgaagaca aggttgtgga cgtggtgcgg 1140
aacctcgtgt tccacctgaa gaagggttac atgattgtca agtgccgggg ccagcaggag 1200
atccaggacc agctgagcct gtccgaagcc ctgcagagag agaagatctt ctttgagaac 1260
caccatatt tcagggatct gctggaggaa ggaaggcca cggttccctg cctggcagaa 1320
aaacttacca gcgagctcat cacacatata tgtaaattct tgccctgtt agaaaatcaa 1380
atcaaggaga ctaccagag aataacagag gagctacaaa agtatgggtg cgacataccg 1440
gaagacgaaa atgaaaaaat gttcttcctg atagataaaa ttaatgcctt taatcaggac 1500
atcactgctc tcactgcaagg agaggaaact gtaggggagg aagacattcg gctgtttacc 1560
agactccgac acgagttcca caaatggagt acaataattg aaaacaattt tcaagaaggc 1620
cataaaattt tgagtagaaa aatccagaaa tttgaaaatc agtatcgtgg tagagagctg 1680
ccaggctttg tgaattacag gacatttgag acaatcgtga aacagcaaat caaggcactg 1740
gaagagccgg ctgtggatat gctacacacc gtgacggata tgggtccggct tgctttcaca 1800
gatgtttcga taaaaaattt tgaagagttt tttaacctcc acagaaccgc caagtccaaa 1860
attgaagaca ttagagcaga acaagagaga gaagggtgaga agctgatccg cctccacttc 1920
cagatggaac agattgtcta ctgccaggac caggatataca ggggtgcatt gcagaaggtc 1980
agagagaagg agctggaaga agaaaaagaa aagaaatcct gggatttttg ggctttccag 2040
tccagctcgg caacagactc ttccatggag gagatctttc agcacctgat ggcctatcac 2100
caggaggcca gcaagcgcac ctccagccac atccctttga tcatccagtt cttcatgctc 2160
cagacgtacg gccagcagct tcagaaggcc atgctgcagc tccctgcagga caaggacacc 2220
tacagctggc tcctgaagga gcggagcgac accagcgaca agcggaagtt cctgaaggag 2280
cggcttgcac ggctgacgca ggctcgccgc cggttgccc agttccccgg ttaaccacac 2340
tctgtccagc cccgtagacg tgcacgcaca ctgtctgccc ccgttccccg gtagccactg 2400
gactgacgac ttgagtgtct agtagtcaga ctggatagtc cgtctctgct tatccgttag 2460
ccgtggtgat ttagcaggaa gctgtgagag cagtttggtt tctagcatga agacagagcc 2520
ccaccctcag atgcacatga gctggcggga ttgaaggatg ctgtcttcgt actgggaaag 2580
ggattttcag cctcagaat cgctccacct tgcagctctc ccttctctct tattcctaga 2640
aactgacaca tgctgaacat cacagcttat ttcctcattt ttataatgtc ccttcacaaa 2700
cccagtgttt taggagcatg agtgccgtgt gtgtgcgtcc tgtcggagcc ctgtctctc 2760
tctctgtaat aaactcattt ctagcag 2787

```

<210> 104

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002462

<400> 104
actgacacat gctgaacatc acagcttatt tcttcatttt tataatgtcc cttcaciaaac 60

<210> 105
<211> 2808
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002759

<400> 105
gcgggcgaggc cgggcgagtt tgctcactt ttgtgacttg cgggtcacagt ggcattcagc 60
tccacacttg gtagaaccac aggcacgaca agcatagaaa catcctaacc aatcttcac 120
gaggcatcga ggtccatccc aataaaaaatc aggagaccct ggctatcata gaccttagtc 180
ttcgctggta tactcgctgt ctgtcaacca gcgggtgact ttttttaagc cttctttttt 240
ctctttttacc agtttctgga gcaaattcag tttgccttcc tggatttgta aattgtaatg 300
acctcaaaac ttttagcagtt ctcccatctg actcaggttt gcttctctgg cggctctcag 360
aatcaacatc cacacttccg tgattatctg cgtgcatttt ggacaaagct tccaaccagg 420
atacggaag aagaaatggc tgggtgatctt tcagcagggt tcttcatgga ggaacttaat 480
acataccgtc agaagcaggg agtagtactt aaatatcaag aactgcctaa ttcaggacct 540
ccacatgata ggagggtttac atttcaagtt ataatagatg gaagagaatt tccagaagg 600
gaaggtagat caaagaagga agcaaaaaat gccgcagcca aattagctgt tgagatactt 660
aataaggaaa agaaggcagt tagtccttta ttattgacaa caacgaattc ttcagaagga 720
ttatccatgg ggaattacat aggcctttatc aatagaattg cccagaagaa aagactaact 780
gtaaatatg aacagtgtgc atcgggggtg catggggccag aaggatttca ttataaatgc 840
aaaatgggac agaaagaata tagtattggg acagggttcta ctaaacagga agcaaaaaca 900
ttggccgcta aacttgcata tcttcagata ttatcagaag aaacctcagt gaaatctgac 960
tacctgtcct ctggttcttt tgctactacg ttgtagtccc aaagcaactc ttttagtgacc 1020
agcacactcg cttctgaatc atcatctgaa ggtgacttct cagcagatac atcagagata 1080
aattctaaca gtgacagttt aaacagttct tcgttgctta tgaatggtct cagaaataat 1140
caaaggaagg caaaaagatc tttggcacc agatttgacc ttcctgacat gaaagaaaca 1200
aagtatactg tggacaagag gtttggcatg gattttaaag aaatagaatt aattgggtca 1260
ggtggatttg gccaggtttt caaagcaaaa cacagaattg acggaagac ttacgttatt 1320
aaacgtgtta aatataataa cgagaaggcg gaggctgaag taaaagcatt ggcaaaactt 1380
gatcatgtaa atattgttca ctacaatggc tgttgggatg gatttgatta tgatcctgag 1440
accagtgatg attctcttga gagcagtgat tatgatcctg agaacagcaa aaatagttca 1500
aggtaaaaga ctaagtgcct tttcatccaa atggaattct gtgataaagg gaccttgga 1560
caatggattg aaaaaagaag aggcgagaaa ctagacaaaag ttttggcttt ggaactcttt 1620
gaacaaataa caaaaggggt ggattatata cattcaaaaa aattaattca tagagatctt 1680
aagccaagta atatatctt agtagatata aaacaagtaa agattggaga ctttggactt 1740
gtaacatctc tgaaaaatga tggaaaagcg acaaggagta agggaacttt gcgatacatg 1800
agccagaac agatttcttc gcaagactat ggaaaggaag tggacctcta cgctttgggg 1860
ctaattcttg ctgaacttct tcatgtatgt gacactgctt ttgaaacatc aaagtttttc 1920
acagacctac gggatggcat catctcagat atatttgata aaaaagaaaa aactotttcta 1980
cagaaattac tctcaaaaga acctgaggat cgacctaaac catctgaaat actaaggacc 2040
ttgactgtgt ggaagaaaag cccagagaaa aatgaacgac acacatgtta gagccttct 2100
gaaaaagtat cctgcttctg atatgcagtt ttccttaaat tatctaaaat ctgctaggga 2160
atatcaatag atatttacct tttatttttaa tgtttccttt aattttttac tatttttact 2220
aatctttctg cagaaacaga aagggttttct tctttttgct tcaaaaacat tcttacattt 2280
tactttttcc tggctcatct ctttattctt tttttttttt ttaaagacag agtctcgctc 2340
tggtgcccag gctggagtgc aatgacacag tcttggctca ctgcaacttc tgctcttgg 2400
gttcaagtga ttctcctgcc tcagcctcct gagtagctgg attacaggca tgtgccacc 2460
acccaactaa tttttgtgtt ttaataaaag acagggtttc accatgttgg ccaggctgg 2520
ctcaaaactc tgacctcaag taatccacct gcctggcct cccaaagtgc tgggattaca 2580
gggatgagcc accgcgccca gcctcatctc tttgttctaa agatggaaaa accaccccca 2640
aattttcttt ttatactatt aatgaatcaa tcaattcata tctatttatt aaatttctac 2700
cgcttttagg ccaaaaaaat gtaagatcgt tctctgcctc acatagctta caagccagct 2760
ggagaaatat ggtactcatt aaaaaaaaaa aaaaagtgat gtacaacc 2808

<210> 106
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002759

<400> 106
 tcgtttctctg cctcacatag cttacaagcc agctggagaa atatggtact cattaaaaaa 60

<210> 107
 <211> 1678
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002811

<400> 107
 aagaaggagg ccgcgcgagg gctgacgaac cggaagaaga ggaactgggc ctgaaaggggt 60
 accggtgacc gctactgctg ccggtgtttg cgtgtggcag ggagccaggc ctggcgagcg 120
 ggggtgtgtcg cgatgccgga gctggcagtg cagaagggtg tgggccaccc cctgggtgctg 180
 ctcagtgtgg tggatcattt caaccgaatc ggcaagggtg gaaaccagaa gcgtgtgtgt 240
 ggtgtgtcttt tgggggtcatg gcaaaagaaa gtacttgatg tatcgaacag ttttgcagtt 300
 ccttttgatg aagatgacaa agacgattct gtatgggttt tagaccatga ttatttggaa 360
 aacatgtatg gaatgtttta gaaagtcaat gccagggaaa gaatagttgg ctggtaccac 420
 acaggcccta aactacacaa gaatgacatt gccatcaacg aactcatgaa aagatactgt 480
 cctaattccg tattggtcat cattgatgtg aagccgaagg acctagggct gcctacagaa 540
 gcgtacattt cagtgaaga agtccatgat gatggaactc caacctcgaa aacatttgaa 600
 cacgtgacca gtgaaattgg agcagaggaa gctgaggaag ttggagttga acacttgta 660
 cgagatatca aagacacgac ggtgggcact ctgtcccagc ggatcacaaa ccagggtccat 720
 ggtttgaagg gactgaactc caagcttctg gatatcagga gctacctgga aaaagtcgcc 780
 acaggcaagc tgcccatcaa ccaccagatc atctaccagc tgcaggacgt cttcaacctg 840
 ctgccagatg tcagcctgca ggagttcgtc aaggcccttt acctgaagac caatgaccag 900
 atgggtgtag tgtacttggc ctgcgtgatc cgttccgtgg tcgccctgca caacctcatc 960
 aacaacaaga ttgccaaccg ggatgcagag aagaaagaag ggcaggagaa agaagagagc 1020
 aaaaaggata ggaaagagga caaggagaaa gataaagata aggaaaagag tgatgtaaag 1080
 aaagaggaga aaaaggagaa aaagtaaaac atgtattaaa tagctttttt aatttgtaaa 1140
 ttaaaatctt acaactaaa tcagtgtgct gctagagggg tctttttcac ttgacatgct 1200
 tattagaaag ctgacccaac aagagctctc tgcctccggt cactcttctg gtgggtgctac 1260
 gtggaagtga atggagactg atctcaaact tgaactgcag ctttcgctgc tgtgagttgg 1320
 ggatatgata gtcagctcag gcttcagatt gtatgagaaa aatgaagaga agtcaacaaa 1380
 tatttttgta ctcttcattc atttatctct aaaaaccagga gttgaatttt cctcatcttg 1440
 aaagactctt ggggtctgtt tctggtattt tacaaaattg ctaagtggaa tgcataaatt 1500
 gcattatgtt ctctggtaac acgtagagtt cagacccttc tgaactctgt tgataatacc 1560
 acaccatgtt ctggacccat agctctggca tcctcagggg ttgtgatcca gctccatata 1620
 ttgtttacct tcaagatac aattaaatgg cttgattttt aaaaaaaaaa aaaaaaaaaa 1678

<210> 108
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002811

<400> 108
 aaattgctaa gtggaatgca tgaattgcat tatgttctct ggtaaacacgt agagttcaga 60

<210> 109

<211> 846
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002888

<400> 109
 ccacgtccgg ggtgccgagc caactttcct gcgtccatgc agccccgcgc gcaacggctg 60
 cccgctccct ggtccggggc cagggggccc cgccccaccg ccccgctgct cgcgctgctg 120
 ctgttgctcg ccccggtggc ggcgcccgcg ggggtccggg gccccgacga ccctgggcag 180
 cctcaggatg ctgggggtccc gcgcaggctc ctgcagcaga aggcgcgcgc ggcgcttcac 240
 ttcttcaact tccgggtccg ctgcgccagc gcgctgcgag tgctggccga ggtgcaggag 300
 ggccgcgcgt ggattaatcc aaaagaggga tgtaaagtcc acgtggctct cagcacagag 360
 cgctacaacc cagagtcttt acttcaggaa ggtgaggac gtttggggaa atgttctgct 420
 cgagtgtttt tcaagaatca gaaaccaga ccaaccatca atgtaacttg tacacggctc 480
 atcgagaaaa agaaaagaca acaaggagat tacctgcttt acaagcaa at gaagcaactg 540
 aaaaacccct tggaaatagt cagcatacct gataatcatg gacatattga tccctctctg 600
 agactcatct gggatttggc ttctcttgga agctcttacg tgatgtggga aatgacaaca 660
 caggtgtcac actactactt ggcacagctc actagtgtga ggcagtgggt aagaaaaacc 720
 tgaaaattaa cttgtgccac aagagttaca atcaaagtgg tctccttaga ctgaattcat 780
 gtgaacttct aatttcatat caagagttgt aatcacattt atttcaataa atatgtgagt 840
 tcctgc 846

<210> 110
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002888

<400> 110
 aaagaaaaga caacaagagg attacctgct ttacaagcaa atgaagcaac tgaaaaaccc 60

<210> 111
 <211> 1054
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003090

<400> 111
 gaattccgcg ggaggccacg ggctttccac agcgcggggg aacgggagggc tgcaggatgg 60
 tcaagctgac ggcgagctg atcgagcagg cggcgcagta caccaacgcg gtgcgcgacc 120
 gggagctgga cctccggggg tataaaattc cgtcattga aaatctaggt gctacgttag 180
 accagtttga tgctattgat tttctgaca atgagatcag gaaactggat ggttttctct 240
 tgttgagaag actgaaaaca ttgttagtga acaacaacag aatatgccgt ataggtgagg 300
 gacttgatca ggctctgccc tgtctgacag aactcattct caccaataat agtctcgtgg 360
 aactgggtga tctggaccct ctggcatctc tcaaactcgt gacttaccta agtatcctaa 420
 gaaatccggt aaccaataag aagcattaca gattgtatgt gatttataaa gttccgcaag 480
 tcagagtact ggatttccag aaagtgaac taaaagagcg tcagggaagca gagaaaatgt 540
 tcaagggcaa acgggggtgca cagcttgcaa aggatattgc caggagaagc aaaactttta 600
 atccagggtc tggtttgcca actgacaaaa agagaggtgg gccatctcca ggggatgtag 660
 aagcaatcaa gaatgccata gcaaatgctt caactctggc tgaagtggag aggctgaagg 720
 ggttgctgca gtctggctcag atccctggca gagaacgcag atcaggggcc actgatgatg 780
 gtgaagaaga gatggaagaa gacacagtca caaacgggtc ctgagcagtg aggcagatgt 840
 ataataatag gccctcttgg aacaagtctt gcttttcgaa catggtataa tagccttgtt 900

tgtgttagca aagtggaatc tatcagcatt gttgaaatgc ttaagactgc tgctgataat 960
 tttgtaatat aagttttgaa atctaaatgt caatttttcta caaattataa aaataaactc 1020
 cactctctat gctaaaaaaa aaaaaaagga attc 1054

<210> 112
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003090

<400> 112
 taatagcctt gtttgtgtta gcaaagtgga atctatcagc attgttgaaa tgcttaagac 60

<210> 113
 <211> 2033
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003158

<400> 113
 gaattccggg actgagctct tgaagacttg ggtccttggt cgcaggtgga ggcacgggtc 60
 tcactccatt gccaggcca gagtgcggga tttttgataa gaaacttcag tgaaggccgg 120
 gcgcggtgct catgcccgta atcccagcat tttcggaggc cgaggcatca tggaccgatc 180
 taaagaaaac tgcatttcag gaacctgttaa ggctacagct ccagttggag gtccaaaacg 240
 tgttctcgtg actcagcaat ttcccttgta gaatccatta cctgtaaata gtggccaggc 300
 tcagcgggtc ttgtgtcctt caaatctctc ccagcgcgtt cctttgcaag cacaaaagct 360
 tgtctccagt cacaagccgg ttocagaatca gaagcagaag caattgcagg caaccagtgt 420
 acctcatcct gtctccaggc cactgaataa cacccaaaag agcaagcagc ccctgccatc 480
 gcacctgaaa ataactcctga ggaggaaactg gcatcaaaac agaaaaatga agaatacaaa 540
 agaggcagtg gcttttggaag acttttgaaat tggctgcctt ctgggttaaag gaaagtttgg 600
 taatgtttat ttggcaagag aaaagcaaaag caagtttatt ctgggtctta aagtgttatt 660
 taaagctcag ctggagaaaag ccggagtgga gcatcagctc agaagagaag tagaaataca 720
 gtcccacctt cggcatccta atattcttag actgtatggt tatttccatg atgtaccag 780
 agtctacctt attctggaat atgcaccact tgggaacagtt tatagagaaac ttcagaaact 840
 ttcaaagttt gatgagcaga gaactgctaa cttatataac agaattgcaa atgccctgtc 900
 ttactgtcat tcgaagagag ttattcatag agacattaag ccagagaact tacttcttgg 960
 atcagctgga gagcttaaaa ttgcagatatt tgggtggtca gtacatgctc catcttccag 1020
 gaggaccact ctctgtggca ccctggacta cctgccccct gaaatgattg aaggctggat 1080
 gcatgatgag aaggtggatc tctggagcct tggagtctct tgctatgaat ttttagttgg 1140
 gaagcctcct tttgaggcaa acacatacca agagacctac aaaagaatat cacgggttga 1200
 attcacattc cctgactttg taacagagggg agccagggac ctcatctcaa gactgttgaa 1260
 gcataatccc agccagaggc caatgctcag agaagtactt gaacacccct ggatcacagc 1320
 aaattcatca aaaccatcaa attgccaaaa caaagaatca gctagcaaac agtcttagga 1380
 atcgtgcagg gggagaaaac cttgagccag ggctgccata taacctgaca ggaacatgct 1440
 actgaagttt attttaccat tgactgctgc cctcaatcta gaacgctaca caagaaatat 1500
 tttgttttta ctacagcagg gtgccttaac ctccctattc agaaagctcc acatcaataa 1560
 acatgacact ctgaagtga agtagccacg agaattgtgc tacttatact ggaacataat 1620
 ctggaggcaa ggctcgaact cagtcgaacc ttgcctccag attatgaacc agtataagta 1680
 gcacaattct cgtggctact ttcacttcag agtgtcatgt ttattgatgt ggagctttct 1740
 gaatagggag gttaaggcac acctgctgag taaaacaaat atttcttgtg tagcgttctt 1800
 aggaatctgg tgtctgtccg gccccgtag gcctgttggg tttctagtcc tcttaccat 1860
 catctccata tgagagtgtg aaaaataggaa cactgtctct acctccattt agggatttgc 1920
 ttgggataca gaagaggcca tgtgtctcag agctgttaag ggcttatttt tttaaaacat 1980
 tggagtcata gcatgtgtgt aaactttaaa tatgcaggcc ttcgtggctc gag 2033

<210> 114
 <211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_003158

<400> 114
ttgggtttct agtcctcctt accatcatct ccatatgaga gtgtgaaaat aggaacacgt 60

<210> 115
<211> 1421
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003258

<400> 115
acttactgcg ggacggcctt ggagagtact cgggttcgtg aacttcccgg aggcgcaatg 60
agctgcatta acctgcccac tgtgtctgcc ggctccccca gcaagacccg ggggcagatc 120
caggtgattc tcgggccgat gttctcagga aaaagcacag agttgatgag acgcgtccgt 180
cgcttcacaga ttgtcagta caagtgcctg gtgatcaagt atgccaaaga cactcgctac 240
agcagcagct totgcacaca tgaccggaac accatggagg cgctgccgc ctgcctgctc 300
cgagaccgtg ccaggaggc cctgggcgtg gctgtcatag gcatcgacga ggggcagttt 360
ttccctgaca tcatggagt ctgcgaggcc atggccaacg ccgggaagac cgtaatgtg 420
gctgcactgg atgggacctt ccagggaag ccatttgggg ccatacctgaa cctgggtgccg 480
ctggccgaga gcgtggtgaa gctgacggcg gtgtgcatgg agtgcttccg ggaagccgcc 540
tataccaaga ggctcggcac agagaaggag gtcgaggtga ttgggggagc agacaagtac 600
cactccgtgt gtcggctctg ctacttcaag aaggcctcag gccagcctgc cgggcccggac 660
aacaagaga actgcccagt gccaggaaaag ccagggggaag ccgtggctgc caggaagctc 720
tttgcccac agcagattct gcaatgcagc cctgccaaact gagggacctg caagggccgc 780
ccgctccctt cctgccactg ccgcctactg gacgctgccc tgcattgctgc ccagccactc 840
caggaggaag tcgggaggcg tggagggtga ccacaccttg gccttctggg aactctcctt 900
tgtgtggctg cccacacctg cgcattgctc ctctctcctt acccaactgg ctgcttaaag 960
cttccctctc agctgctggg acgatcgccc aggcctggagc tggccccgct tgggtggcctg 1020
ggatctggca cactccctct ccttgggggtg agggacagag cccacgctg ttgacatcag 1080
cctgcttctt ccctctgctg gctttcactg ctgagtttct gttctccctg ggaagcctgt 1140
gccagcacct ttgagccttg gccacactg aggccttaggc ctctctgcct gggatgggct 1200
cccaccctcc cctgaggatg gcctggattc acgcctctt gtttcccttt gggctcaaag 1260
cccttcctac ctctggtgat ggtttccaca ggaacaacag catctttcac caagatgggt 1320
ggcaccaacc ttgctgggac ttggatccca ggggcttatc tcttcaagtg tggagagggc 1380
agggtccacg cctctgctgt agcttatgaa attaactaat t 1421

<210> 116
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003258

<400> 116
cttcctacct ctgggtgatgg tttccacagg aacaacagca tctttcacca agatgggtgg 60

<210> 117
<211> 913
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003311

<400> 117

```

agagccggcg ccgtcaccgc ccgcattgcc gctcccagtc ccgcgctcgg caccacatga 60
aatcccccca cgaggtgcta ccgcaggggcg agttggagaa gcgcagcgac agcctcttcc 120
agctatggaa gaagaagcgc ggggtgctca cctccgaccg cctgagcctg ttccccgcca 180
gcccccgccg gcgccccaaag gagctgcgct tccactccat cctcaagggtg gactgcgtgg 240
agcgcacggg caagtacgtg tacttcacca tctgcaccac cgaccacaag gagatcgact 300
tccgctgcgc gggcgagagc tcttggaacg cggccatcgc gctggcgctc atcgatttcc 360
agaaccgccg cgccctgcag gactttcgca gccgccagga acgcaccgca cccgccgcac 420
ccgccgagga cgccgtggct gccgcggccg ccgcaccctc cgagccctcg gagccctcca 480
ggccatcccc gcagcccaaa ccccgcacgc catgagcccc ccgcgggcca tacgctggac 540

```

```

gagtcggacc gaggttagga cgtggccggc gctctccagc cctgcagcag aagaacttcc 600
cgtgcgcgcg gatcctcgtc ccgttgcaac ggcgccttaa gttattggac tatctaatat 660
ctatgtatatt atttcgctgg ttctttgtag tcacatatatt tatagtctta atatcttggt 720
tttgcatcac tgtgcccatt gcaaataaat cacttgggca gtttgctttt ctaccatccg 780
gctgtggctc agtgagactc ctgctgggag ggtggaggcc cagggaatggg cgggcaggac 840
accctcatcc agtcctgcgg ggtcgtgtg aaaggcgctg ggaaccggct ttgaatgaat 900
aatgaatcg tgt 913

```

<210> 118

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003311

<400> 118

```

atttcgctgg ttctttgtag tcacatatatt tatagtctta atatcttggt ttgcatcac 60

```

<210> 119

<211> 1723

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003376

<400> 119

```

tcgcggaggc ttggggcagc cgggtagctc ggagggtcgtg gcgctggggg ctagcaccag 60
cgctctgtcg ggaggcgagc cgggttaggtg gaccgggtcag cggactcacc ggccaggcg 120
ctcgggtgctg gaatttgata ttcattgata cgggttttat cctcttctt ttttcttaaa 180
catttttttt taaaactgta ttgtttctcg ttttaattta tttttgcttg ccattcccca 240
cttgaatcgg gccgacggct tggggagatt gctctacttc cccaaatcac tgtggatttt 300
ggaaaccagc agaaagagga aagaggtagc aagagctcca gagagaagtc gaggaagaga 360
gagacggggt cagagagagc gcgcgggctg gcgagcagcg aaagcgacag gggcaaagtg 420
agtgcctgct ttttgggggt gaccgcggga gcgcggcgct agccctcccc cttgggatcc 480
cgcagctgac cagtgcgct gacggacaga cagacagaca ccgccccag cccagctac 540
cacctcctcc ccggccgctg gcggacagtg gacgcggcgg cgagccgcgg gcaggggccg 600
gagcccgctc ccggaggcgg ggtggagggg gtcggggctc gcggcgctgc actgaaactt 660
ttcgtccaac ttctgggctg ttctcgcttc ggaggagccg tgggtccgctc gggggaagcc 720
gagccgagcg gagccgcgag aagtgtctagc tcgggcccgg aggagccgca gccggaggag 780
ggggaggagg aagaagagaa ggaagaggag agggggccgc agtggcgact cggcgctcgg 840
aagccgggct catggacggg tgaggcggcg gtgtgcgcag acagtgtctc agccgcgcgc 900
gctccccagg ccttgcccg ggccctcggg cggggaggaa gactagctcg ccgaggcgcc 960
gaggagagcg ggccgcccc cagcccgagc cgggagggga gcgcgagccg ccggggcccc 1020
ggtcgggcct ccgaaaccat gaactttctg ctgtcttggg tgcattggag ccttgccctt 1080
ctgctctacc tccaccatgc caagtggctc caggctgcac ccatggcaga aggaggaggg 1140
cagaatcatc acgaagtggg gaagtccatg gatgtctatc agcgcagcta ctgccatcca 1200
atcgagaccc tgggtggacat cttccaggag taccctgatg agatcgagta catcttcaag 1260

```

```

ccatcctgtg tgccccctgat gcgatgcggg ggctgctgca atgacgaggg cctggagtgt 1320
gtgcccactg aggagtccaa catcaccatg cagattatgc ggatcaaacc tcaccaaggc 1380
cagcacatag gagagatgag ctccctacag cacaacaaat gtgaatgcag accaaagaaa 1440
gatatagcaa gacaagaaaa aaaatcagtt cgaggaaagg gaaaggggca aaaacgaaag 1500
cgcaagaaat cccgggtataa gtcctggagc gttccctgtg ggccttgctc agagcggaga 1560
aagcatttgt ttgtacaaga tccgcagacg tgtaaatgtt cctgcaaaaa cacagactcg 1620
cgttgcaagg cgaggcagct tgagttaaac gaacgtactt gcagatgtga caagccgagg 1680
cggtgagccg ggcaggagga aggagcctcc ctcagggttt cgg 1723

```

<210> 120
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003376

<400> 120
 ccagcacata ggagagatga gcttcctaca gcacaacaaa tgtgaatgca gaccaaagaa 60

<210> 121
 <211> 2834
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003406

<400> 121
 gccactccc accgccagct ggaaccctgg ggactacgac gtccctcaaa ccttgctttct 60
 aggagataaa aagaacatcc agtcatggat aaaaatgagc tggttcagaa ggccaaactg 120
 gccgagcagg ctgagcgata tgatgacatg gcagcctgca tgaagtctgt aactgagcaa 180
 ggagctgaat tatccaatga ggagaggaat cttctctcag ttgcttataa aaatgtttgta 240
 ggagcccgtg ggtcatcttg gagggctcgtc tcaagtattg aacaaaagac ggaaggtgct 300
 gagaaaaaac agcagatggc tcgagaatac agagagaaaa ttgagacgga gctaagagat 360
 atctgcaatg atgtactgtc tcttttggaa aagtctctga tccccaatgc ttcacaagca 420
 gagagcaaag tcttctatct gaaaatgaaa ggagattact accgttactt ggctgaggtt 480
 gccgctgggtg atgacaagaa agggattgtc gatcagtcac aacaagcata ccaagaagct 540
 tttgaaatca gcaaaaagga aatgcaacca acacatccta tcagactggg tctggccctt 600
 aacttctctg tgttctatta tgagattctg aactccccag agaaagcctg ctctcttgca 660
 aagacagctt ttgatgaagc cattgctgaa cttgatcacat taagtgaaga gtcatacaaa 720
 gacagcacgc taataatgca attactgaga gacaacttga cattgtggac atcggatacc 780
 caaggagacg aagctgaagc aggagaagga ggggaaaatt aaccggcctt ccaacttttg 840
 tctgcctcat tctaaaatct acacagtaga ccatttgtca tccatgctgt cccacaaata 900
 gtttttttgt taccgatttat gacaggttta tgttacttct atttgaattt ctatatcttc 960
 catgtggttt ttatgtttta tattagggga gtagagccag ttaacattta gggagttatc 1020
 tgttttctc tttgaggtggc caatatgggg atgtggaatt tttatacaag ttataagtgt 1080
 ttggcatagt acttttggt cattgtggct tcaaaaagggc cagtgtaaaa ctgcttccat 1140
 gtctaagcaa agaaaactgc ctacatactg gtttgcctg gcggggaata aaagggatca 1200
 ttggttccag tcacaggtgt agtaattgtg ggtactttta gggttgagc acttacaagg 1260
 ctgtggtaga atcatacccc atggatacca catattaaac catgtatatc tgtggaatac 1320
 tcaatgtgta cacttttgac tacagctgca gaagtgttcc tttagacaaa gttgtgaccc 1380
 attttactct ggataagggc agaaacgggt cacattccat tatttgtaaa gttacctgct 1440
 gttagctttc attatttttg ctacactcat tttatttgta tttaaatgtt ttaggcaacc 1500
 taagaacaaa tgtaaaagta aagatgcagg aaaaatgaat tgcttggtat tcattacttc 1560
 atgtatatca agcacagcag taaaacaaaa acccatgtat ttaacttttt tttaggattt 1620
 ttgcttttgt gatTTTTTTT tttttttttt gatacttgcc taacatgcat gtgctgtaaa 1680
 aatagttaac agggaaataa cttgagatga tggctagctt tgtttaatgt cttatgaaat 1740
 tttcatgaac aatccaagca taattgttaa gaacacgtgt attaaattca tgtaagtgga 1800
 ataaaagttt tatgaatgga cttttcaact actttctcta cagcttttca tgtaaatagg 1860

```

tcttggttct gaaacttctc taaaggaaat tgtacattht ttgaaattta ttccttatte 1920
cctcttggca gctaattggc tcttaccaag tttaaacaca aaatttatca taacaaaaat 1980
actactaata taactactgt ttccatgtcc catgatcccc tctcttcctc cccaccctga 2040
aaaaaatgag ttcttatttt ttctgggaga gggggggatt gattagaaaa aaatgtagtg 2100
tgttccattt aaaatttttg catatggcat tttctaactt aggaagccac aatgttcttg 2160
gcccacatg acattgggta gcattaactg taagttttgt gcttccaaat cacttttttg 2220
tttttaagaa tttcttgata ctcttatagc ctgccttcaa ttttgatcct ttattctttc 2280
tatttgtcag gtgcacaaga ttaccttcct gttttagcct tctgtcttgt caccaacct 2340
tcttacttg tggccatgta cttggaaaaa ggccgcatga tctttctggc tccactcagt 2400
gtctaaggca ccctgcttcc tttgcttgca tcccacagac tatttcctc atcctattta 2460
ctgcagcaaa tctctcctta gttgatgaga ctgtgtttat ctccctttta aacctacct 2520
atcctgaatg gtctgtcatt gtctgccttt aaaatcctc ctctttcttc ctctctatt 2580
ctctaaataa tgatggggct aagttatacc caaagctcac tttaaaaaat atttcctcag 2640
tactttgcag aaaacaccaa acaaaaatgc cattttaaaa aagggtgtatt ttttctttta 2700
gaatgtaagc tcctcaagag caggggacaat gttttctgta tgttctattg tgcctagtac 2760

actgtaaatg ctcaataaat attgatgatg ggaggcagtg agtcttgatg ataaggggtga 2820
gaaactgaaa tccc 2834

```

<210> 122

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003406

<400> 122

tttagccttc tgtcttgtca ccaaccattc ttacttgggtg gccatgtact tggaaaaagg 60

<210> 123

<211> 1938

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003504

<400> 123

```

gatttggcgg gagtcttgac cgcgcgcggg ctcttgggtac ctcagcgcga gcgccaggcg 60
tccggccgcc gtggctatgt tctgttccga ttccgcgaaa gatttctacg aggtgggtcca 120
gagccagagg gtctctctct tctgtggcctc ggaactggat gctctgtgtg cgtgcaagat 180
ccttcaggcc ttgttccagt gtgaccacgt gcaatatacg ctggttccag tttctgggtg 240
gcaagaactt gaaactgcat ttcttgagca taaagaacag ttctattatt ttattctcat 300
aaactgtgga gctaattgtg acctattgga tattcttcaa cctgatgaag acactatatt 360
ctttgtgtgt gacacccata ggccagtcaa tgtcgtcaat gtatacaacg ataccagat 420
caaattactc attaaacaag atgatgacct tgaagttccc gcctatgaag acatcttcag 480
ggatgaagag gaggatgaag agcattcagg aaatgacagt gatgggtcag agccttctga 540
gaagcgcaac cggtagaag aggagatagt ggagcaaac atgcggagga ggcagcggcg 600
agagtgggag gcccgagaa gagacatcct ctttgactac gagcagtag aatatcatgg 660
gacatcgtca gccatggtga tgtttgagct ggcttggatg ctgtccaagg acctgaatga 720
catgctgtgg tgggccatcg ttggactaac agaccagtgg gtgcaagaca agatcactca 780
aatgaaatac gtgactgatg ttggtgtcct gcagcgcac gtttcccgcg acaaccaccg 840
gaacgaggat gaggagaaca cactctccgt ggactgcaca cggatctcct ttgagtatga 900
cctccgctg gtgctctacc agcactgggtc cctccatgac agcctgtgca acaccagcta 960
taccgcagcc aggttcaagc tgtggtctgt gcatggacag aagcggctcc aggtttcct 1020
tgcagacatg ggtcttcccc tgaagcaggt gaagcagaag ttccaggcca tggacatctc 1080
cttgaaggag aatttgccgg aaatgattga agagtctgca aataaatttg ggatgaagga 1140
catgcgcgtg cagactttca gcattcattt tgggttcaag cacaagtttc tggccagcga 1200
cgtggtcttt gccaccatgt ctttgatgga gagccccgag aaggatggct caggagacaga 1260
tcacttcac caggctctgg acagcctctc caggagtaac ctggacaagc tgtaccatgg 1320

```

```

cctggaactc gccagaagc agctgagagc caccagcag accattgcc gctgcctttg 1380
caccaacctc gtcattctcc aggggccttt cctgtactgc tctctcatgg agggcactcc 1440
agatgtcatg ctgtttctta ggcgggcac cctaagcctg ctgagcaaac acctgctcaa 1500
gtcctttgtg tggttcgaca agaaccggcg ctgcaaacctg ctgcccctgg tgatggctgc 1560
ccccctgagc atggagcatg gcacagtgc cgtggtgggc atccccccag agaccgacag 1620
ctcggacagg aagaactttt ttgggagggc gtttgagaag gcagcggaaa gcaccagctc 1680
ccggatgctg cacaaccatt ttgacctctc agtaattgag ctgaaagctg aggatcggag 1740
caagtttctg gacgcactta tttccctcct gtccataggaa tttgattctt ccagaatgac 1800
cttcttattt atgtaactgg ctttcattta gattgtaagt tatggacatg atttgagatg 1860
tagaagccat tttttattaa ataaaatgct tatttttaggc tccgtcccca aaaaaaaaaa 1920
aaaaaaaaaa aaaaaaaaaa 1938

```

<210> 124

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003504

<400> 124

```

caagtttctg gacgcactta tttccctcct gtccataggaa tttgattctt ccagaatgac 60

```

<210> 125

<211> 2346

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003600

<400> 125

```

acaaggcagc ctgcgtcgag cgcaggccaa tgggctttct agctagaggg ttttaactcct 60
atttataaaag aagaaccttt gaattctaac ggctgagctc ttggaagact tgggtccttg 120
ggctcgcagggt gggagccgac ggggtgggtag accgtggggg atatctcagt ggcggacgag 180
gacggcggggg acaagggggcg gctggtcgga gtggcggagc gtcaagtccc ctgtcggttc 240
ctcgttccct gagtgtcctt ggcgtgcctt tgtcccgc cagcgccttt gcatccgttc 300
ctgggcacccg aggcgcctct taggatactg cttgttactt attacagcta gaggcatcat 360
ggaccgatct aaagaaaact gcatttcagg acctgttaag gctacagctc cagttggagg 420
tccaaaacgt gttctcgtga ctacgaatt tccctgtcag aatccattac ctgtaaatag 480
tggccaggct cagcgggtct tgtgtccttc aaattcttcc cagcgcattc ctttgcaagg 540
acaaaagctt gtctccagtc acaagccggt tcagaatcag aagcagaagc aattgcaggc 600
aaccagtgtg cctcatcctg tctccaggcc actgaataac acccaaaaga gcaagcagcc 660
cctgccatcg gcacctgaaa ataactctga ggaggaactg gcatcaaaac agaaaaatga 720
agaatcaaaa aagaggcagt gggctttgga agactttgaa attggtcgcc ctctgggtaa 780
aggaaagttt ggtaatgttt atttggcaag agaaaagcaa agcaagttta ttctggctct 840
taaagtgtta tttaaagctc agctggagaa agccggagtg gagcatcagc tcagaagaga 900
agtagaaata cagtcaccac ttcggcatcc taatattctt agactgtatg gttatttcca 960
tgatgctacc agagtctacc taattctgga atatgcacca cttggaacag tttatagaga 1020
acttcagaaa ctttcaaagt ttgatgagca gagaactgct acttatataa cagaattggc 1080
aatgcctctg tcttactgtc attcgaagag agttattcat agagacatta agccagagaa 1140
cttacttctt ggatcagctg gagagcttaa aattgcagat ttgggtgggt cagtacatgc 1200
tccatcttcc aggaggacca ctctctgtgg caccctggac tacctgcccc ctgaaatgat 1260
tgaaggtcgg atgcatgatg agaaggtgga tctctggagc cttggagttc tttgctatga 1320
atthtttagtt gggaagcttc cttttgaggc aaacacatac caagagacct acaaaaagaat 1380
atcacgggtt gaattcacat tccctgactt tgtaacagag ggagccaggg acctcatttc 1440
aagactgttg aagcataatc ccagccagag gccaatgctc agagaagtac ttgaacaccc 1500
ctggatcaca gcaaatctat caaaaccatc aaattgcaa aacaaagaat cagctagcaa 1560
acagtcttag gaatcgtgca gggggagaaa tccttgagcc agggctgcca tataacctga 1620
caggaacatg ctactgaagt ttattttacc attgactgct gccctcaatc tagaacgcta 1680
cacaagaaat atttgtttta ctcagcagggt gtgccttaac ctccctattc agaaagctcc 1740

```

```

acatcaataa acatgacact ctgaagtga agtagccacg agaattgtgc tacttatact 1800
ggttcataat ctggaggcaa ggttcgactg cagccgcccc gtcagcctgt gctaggcatg 1860
gtgtcttcac aggaggcaaa tccagagcct ggctgtgggg aaagtgacca ctctgccctg 1920
accccgatca gtttaaggagc tgtgcaataa ccttcctagt acctgagtga gtgtgtaact 1980
tattgggttg gcgaagcctg gtaaagctgt tggaaatgagt atgtgattct ttttaagtat 2040
gaaaataaag atatatgtac agacttgtat tttttctctg gtggcattcc ttttaggaatg 2100
ctgtgtgtct gtccggcacc ccggtaggcc tgattggggt tctagtcctc cttaaccact 2160
tatctcccat atgagagtgt gaaaaatagg aacacgtgct ctacctccat ttagggattt 2220
gcttgggata cagaagaggc catgtgtctc agagctgtta agggcttatt tttttaaacc 2280
attggagtca tagcatgtgt gtaaacctta aatatgcaa taaataagta tctatgtcta 2340
aaaaaa 2346

```

```

<210> 126
<211> 60
<212> DNA
<213> Homo sapiens

```

```
<300>
```

```
<308> NM_003600
```

```

<400> 126
agagtgtgaa aaataggaac acgtgctcta cctccattta gggatttgct tgggatacag 60

```

```

<210> 127
<211> 853
<212> DNA
<213> Homo sapiens

```

```
<300>
```

```
<308> NM_003641
```

```

<400> 127
ctagtccctga cttcactttct gatgaggaag cctctctcct tagccttcag cctttcctcc 60
caccctgcca taagtaattt gatcctcaag aagttaaacc acacctcatt ggteccctggc 120
taattcacca atttacaac agcaggaaat agaaacttaa gagaaataca cacttctgag 180
aaactgaaac gacaggggaa aggaggtctc actgagcacc gtcccagcat ccggacacca 240
cagcggccct tcgctccacg cagaaaacca cacttctcaa accttccctc aacacttctc 300
tccccaaagc cagaagatgc acaaggagga acatgaggtg gctgtgctgg gggcaccctc 360
cagcaccatc cttccaaggt ccaccgtgat caacatccac agcgagacct ccgtgccccga 420
ccatgtcgtc tggteccctgt tcaacaccct cttcttgaac tgggtgctgtc tgggcttcat 480
agcattcgcc tactccgtga agtctaggga caggaagatg gttggcgacg tgaccggggc 540
ccaggccctat gcctccaccg ccaagtgcct gaacatctgg gccctgattc tgggcatcct 600
catgaccatt ggattcatcc tgtcactggg attcggctct gtgacagtct accatattat 660
gttacagata atacaggaaa aacgggggta ctagtaccg cccatagcct gcaacctttg 720
cactccactg tgcaatgctg gccctgcacg ctggggctgt tgcccctgcc cccttgggtc 780
tgcccctaga tacagcagtt tataccacac cacctgtcta cagtgtcatt caataaagtg 840
cacgtgcttg tga 853

```

```

<210> 128
<211> 60
<212> DNA
<213> Homo sapiens

```

```
<300>
```

```
<308> NM_003641
```

```

<400> 128
attatgttac agataatata ggaaaaacgg ggttactagt agccgccccat agcctgcaac 60

```

<210> 129
 <211> 1280
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003756

<400> 129
 gaaagatggc gtcccgcaag gaaggtaccg gctctactgc cacctcttcc agctccaccg 60
 ccggcgccagc agggaaaggc aaaggcaaaag gcgggctcggg agattcagcc gtgaagcaag 120
 tgcagataga tggccttgtg gtattaaaga taatcaaca ttatcaagaa gaaggacaag 180
 gaactgaagt tgttcaagga gtgcttttgg gtctggttgt agaagatcgg cttgaaatta 240
 ccaactgctt tcccttccct cagcacacag aggatgatgc tgactttgat gaagtccaat 300
 atcagatgga aatgatgcgg agccttcgcc atgtaaacad tgatcatctt cactggggct 360
 ggtatcagtc cacatactat ggctcattcg ttaccggggc actcctggac tctcagttta 420
 gttaccagca tgccattgaa gaatctgtcg ttctcattta tgatcccata aaaactgccc 480

 aaggatctct ctactaaag gcatacagac tgactcctaa actgatggaa gtttgtaaag 540
 aaaaggattt tccccctgaa gcattgaaaa aagcaaatat cacctttgag tacatgtttg 600
 aagaagtgcc gattgtaatt aaaaattcac atctgatcaa tgtcctaata tgggaacttg 660
 aaaagaagtc agctgtttgca gataaacatg aattgctcag ccttgccagc agcaatcatt 720
 tggggaagaa tctacagttg ctgatggaca gagtggatga aatgagccaa gatatagtta 780
 aatacaacac atacatgagg aatactagta aacaacagca gcagaaacat cagtatcagc 840
 agcgtcgcca gcaggagaat atgcagcgcc agagccgagg agaaccctcg ctccctgagg 900
 aggacctgtc caaactcttc aaaccaccac agccgcctgc caggatggac tgcgtgctca 960
 ttgcaggcca gataaacact tactgccaga acatcaagga gttcactgcc caaaacttag 1020
 gcaagctctt catggcccag gctcttcaag aatacaacaa ctaagaaaag gaagtttcca 1080
 gaaaagaagt taacatgaac tcttgaagtc acaccagggc aactcttgga agaaatatat 1140
 ttgcataattg aaaagcacag aggatttctt tagtgtcatt gccgattttg gctataacag 1200
 tgtctttcta gccataataa aataaaaaaa aaaaaaaa aaaaaaaaaa 1260
 aaaaaaaaaa aaaaaaaaaa 1280

<210> 130
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003756

<400> 130
 tgagccaaga tatagttaaa tacaacacat acatgaggaa tactagtaaa caacagcagc 60

<210> 131
 <211> 839
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003832

<400> 131
 aagccacagg ctccctggct ggcgtcagct aaagtggctg ttgggtgtcc gcaggcttct 60
 gcttggccgc gcgcgcctat aagctaccag gaggagcttt acgacttccc gtccctgcggg 120
 aagtggcggg cagcatcgca aggtagcgca gaagcttctc aatggccagc gccagctgca 180
 gccccggcgg cgactcgcc tcacctgagc ctgggaggaa aattcttcca aggatgatct 240
 cccactcaga gctgaggaag cttttctact cagcagatgc tgtgtgtttt gatgttgaca 300
 gcacggtcat cagtgaagaa ggaatcggat gctttcattg gatttggagg aaatgtgatc 360
 aggcaacaag tcaaggataa cgccaaatgg tatatcactg attttgtaga gctgctggga 420

```

gaaccggaag aataacatcc attgtcatatc agctccaaac aacttcagat gaatTTTTTtac 480
aagttacaca gattgatact gtttgcttac aattgcctat tacaacttgc tataaaaaagt 540
tggtacagat gatctgcact gtcaagtaaa ctacagttag gaatcctcaa agattgggttt 600
gtttgtttttt aactgtagtt ccagtattat atgatcacta tcgatttcoct ggagagtttt 660
gtaatctgaa ttctttatgt atattcctag ctatatTTTca tacaaagtgt ttttaagagtg 720
gagagtcaat taaacacctt tactcttagg aatatagatt cggcagcctt cagtgaatat 780
tggtttttttt cccttttggtg tgtcaataaaa agtttatcca tgtgtcagaa aaaaaaaaa 839

```

<210> 132

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003832

<400> 132

```

gaagaaggaa tcggatgctt tcattggatt tggaggaaat gtgatcaggc aacaagtcaa 60

```

<210> 133

<211> 3128

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003981

<400> 133

```

gcttcgcccc gtggcgcggt ttgaaatTTT gcggggctca acggctcgcg gagcggctac 60
gcggagtgc atcgccggtg tttgcgggtg gttgttgctc tcggggccgt gtggagtagg 120
tctggacctg gactcacggc tgcttgagc gtccgccatg aggagaagtg aggtgctggc 180
ggaggagtcc atagtatgtc tgcagaaagc cctaaatcac ctccgggaaa tatgggagct 240
aattgggatt ccagaggacc agcggttaca aagaactgag gtggtaaaga agcatatcaa 300
ggaactcctg gatatgatga ttgctgaaga ggaaagcctg aaggaaagac tcatcaaaag 360
catatccgtc tgtcagaaag agctgaacac tctgtgcagc gaggttacatg ttgagccatt 420
tcaggaagaa ggagagacga ccatcttgca actagaaaaa gatttgcgca cccaagtggg 480
attgatgcga aaacagaaaa aggagagaaa acaggaaactg aagctacttc aagagcaaga 540
tcaagaactg tgcgaaattc tttgtatgcc ccactatgat attgacagtg cctcagtgcc 600
cagcttagaa gagctgaacc agttcaggca acatgtgaca actttgaggg aaacaaaggc 660
ttctaggcgt gaggagtttg tcagtataaa gagacagatc atactgtgta tggagaagatt 720
agaccacacc ccagacacaa gctttgaaag agatgtgggtg tgtgaagacg aagatgcctt 780
ttgtttgtct ttggagaata ttgcaacact acaaaaagttg ctacggcagc tggaaatgca 840
gaaatcacia aatgaagcag tgtgtgaggg gctgcgtact caaatccgag agctctggga 900
caggttgcaa atacctgaag aagaaagaga agctgtggcc accattatgt ctgggtcaaa 960
ggccaaggtc cggaaagcgc tgcaattaga agtggatcgg ttggaagaac tgaaaatgca 1020
aaacatgaag aaagtgattg aggcatttcg agtggagctg gttcagtaact gggaccagtg 1080
cttttatagc caggagcaga gacaagcttt tgccctttt tgtgtgaggg actacacaga 1140
aagtctgtc cagctccacg atgctgagat tgtgcggtta aaaaactact atgaagttca 1200
caaggaaactc tttgaagggtg tccagaagtg ggaagaaacc tggaggcttt tcttagagtt 1260
tgagagaaaa gcttcagatc caaatcgatt tacaaccga ggaggaaatc ttctaaaaga 1320
agaaaaacaa cgagccaagc tccagaaaat gctgcccaag ctggaagaag agttgaaggc 1380
acgaattgaa ttgtgggaac aggaacattc aaaggcattt atggtgaatg ggcagaaatt 1440
catggagtat gtggcagaac aatgggagat gcatcgattg gagaaagaga gagccaagca 1500
ggaaagacaa ctgaagaaca aaaaacagac agagacagag atgctgtatg gcagcgctcc 1560
tcgaacacct agcaagcggc gaggactggc tccaataca ccgggcaaaag cacgtaagct 1620
gaacactacc accatgtcca atgctacggc caatagtagc attcggccta tctttggagg 1680
gacagtctac cactcccccg tgtctcgact tcctccttct ggcagcaagc cagtcgctgc 1740
ttccacctgt tcagggaaga aaacaccccg tactggcagg catggagcca acaaggagaa 1800
cctggagctc aacggcagca tectgagtgg tgggtaccct ggctcggccc ccctccagcg 1860
caacttcagc attaattctg ttgccagcac ctattctgag tttgcgaagg atccgtccct 1920

```

```

ctctgacagt tccactgttg ggcttcagcg agaactttca aaggcttcca aatctgatgc 1980
tactttctgga atcctcaatt caaccaacat ccagtcctga gaagccctga tcagtcaacc 2040
agctgtggct tcctgtgcct agactggacc taattatatg ggggtgactt tagtttttct 2100
tcagcttagg cgtgcttgaa accttggcca ggttccatga ccatgggcct aacttaaaga 2160

tgtgaatgag tggttacagtt gaaagcccat cataggttta gtggtcctag gagacttggt 2220
tttgacttat atacatgaaa agtttatggc aagaagtgc aatttttagca tatggggcct 2280
gactttctcta ccacataatt ctacttgctg aagcatgatc aaagcttggt ttatttcacc 2340
actgtaggaa aatgattgac tatgcccatc cctgggggta attttggcat gtatacctgt 2400
aactagtaat taacatcttt tttgttttagg catgttcaat taatgctgta gctatcatag 2460
cttttgcctt accctgaagc ttgtcccccac cacacaggac agccttcctc ctgaagagaa 2520
tgtctttgtg tgccgaagt tgagatggcc tgccctactg ccaaagaggt gacaggaagg 2580
ctgggagcag ctttgtttaa ttgtgttcag ttctgttaca cagtgcattg ccctttgttg 2640
ggggtatgca tgtatgaaca cacatgcttg tcggaacgct ttctcggcgt ttgtcccttg 2700
gctctcatct ccccatctcc tgtgcctact ttgcctgagt tcttctaccc ccgcagttgc 2760
cagccacatt gggagttctgt ttgttccaat gggttgagct gtctttgtcg tggagatctg 2820
gaactttgca catgtcacta ctggggaggt gttcctgctc tagcttccac gatgaggcgc 2880
cctctttacc tatcctctca atcactactc ttcttgaagc actattatbt attcttccgc 2940
tgtctgcctg cagcagtaact actgtcaaca tagtgtaaat ggttctcaaa agcttaccag 3000
tgtggacttg gtgttagcca cgctgtttac tcatacagta cgtgtcctgt ttttaaaata 3060
tacaattatt cttaaaaata aattaaaatc tgtatactta catttcaaaa agaaaaaaa 3120
aaaaaaa 3128

```

<210> 134
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003981

```

<400> 134
tgcagcagta ctactgtcaa catagtgtaa atggttctca aaagcttacc agtgtggact 60

```

<210> 135
<211> 1816
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004029

```

<400> 135
ggcaccaggt gtccggcctg cgccttcccg ccaggcctgg aactgggttc aacacctgtg 60
acttcatgtg tgcgcgccgg ccacacctgc agtcacacct gtagcccccct ctgccaagag 120
atccataccg aggcagcgtc ggtggctaca agcctcagt ccacacctgt ggacacctgt 180
gacacctggc cacacgacct gtggccgcgg cctggcgtct gctgcgacag gagcccttac 240
ctccccgttt ataacacctg accgccacct aactgcccct gcagaaggag caatggcctt 300
ggctcctgag agggcagccc cagcgtgtgt gttcggagag tggctccttg gagagatcag 360
cagcggctgc tatgaggggc tgcagtggct ggacgaggcc cgcacctgtt tccgcgtgcc 420
ctggaagcac ttgcgcgcga aggacctgag cgaggccgac gcgcgcctct tcaaggcctg 480
ggctgtggcc cgcggcaggt ggccgcctag cagcagggga ggtggccgc ccccgaggc 540
tgagactgcg gagcgcgccg gctggaaaac caacttccgc tgcgcactgc gcagcacgcg 600
tcgcttcgtg atgctgctggg ataactcggg ggacccggcc gaccgcgaca aggtgtacgc 660
gctcagccgg gagctgtgct ggcgagaagg cccaggcacg gaccagactg aggcagaggc 720
ccccgcagct gtcccaccac cacagggtgg gcccccaggg ccattcttgg cacacacaca 780
tgtgggactc caagccccag gccccctccc tgccccagct ggtgacaagg gggacctcct 840
gctccaggca gtgcaacaga gctgcctggc agaccatctg ctgacagcgt catggggggc 900
agatccagtc ccaaccaagg ctctggaga gggacaagaa gggcttccc tgactggggc 960
ctgtgctgga ggcgaggccg cggccccaga gtccccgcac caggcagagc cgtacctgtc 1020
accctcccca agcgcctgca ccgcggtgca agagcccagc ccaggggcgc tggacgtgac 1080

```



```

catcatgtac aagggccgca cgggtgctgca gaaggtggtg ggacacccga gctgcacgtt 1140
cctatacggc cccccagacc cagctgtccg ggccacagac cccagcagg tagcattccc 1200
cagccctgcc gagctcccg accagaagca gctgcgtac acggaggaac tgctgcggca 1260
cgtggccctt gggttgcacc tggagcttcg ggggcccacag ctgtggggccc ggcgcatggg 1320
caagtgcgaag gtgtactggg aggtggggcg acccccaggc tccgccagcc cctccacccc 1380
agcctgcctg ctgcctcgga actgtgacac ccccatcttc gacttcagag tcttcttcca 1440
agagctggtg gaattccggg cacggcagcg cegtggctcc ccacgctata ccatctacct 1500
gggcttcggg caggacctgt cagctgggag gcccaaggag aagagcctgg tcttggtgaa 1560
gctggaaccc tggctgtgcc gagtgcacct agagggcacg cagcgtgagg gtgtgtcttc 1620
cctggatagc agcagcctca gcctctgcct gtccagcgcc aacagcctct atgacgacat 1680
cgagtgtctt cttatggagc tggagcagcc cgcctagaac ccagtctaata gagaactcca 1740
gaaagctgga gcagcccacc tagagctggc cgcggccgcc cagtctaata aaaagaactc 1800
cagaacaaaa aaaaaa 1816

```

<210> 136
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004029

<400> 136
 agcagccac ctagagctgg ccgcccgc ccagtctaata aaaaagaact ccagaacaaa 60

<210> 137
 <211> 2121
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004203

<400> 137
 tggaattttt ggcgcgagca gctccgcgcg cgttcacggg ccgttcccc tcacgggaggt 60
 cctccgcccg ggctccgga acagtcgacg gcagactccg gcccgctgag ccacccgagg 120
 ggtcccgtgg cctccgcgga cccggaatct gggccctcgc ggacccgcgc cccgcccagt 180
 cgccccaggg cttccccaca cccacggagt gaagtcagcc gcggccctgc ctggggaggaa 240
 cttaccgtct accgggaaag gtggccagca gatgtgtcgg gcctggtgag aggggtgaggc 300
 gagacggccc gatcgcccag ggccccggaa gctgcggagg tcacccccgc ctggccttag 360
 ctcagggaca ccctggattc acgtgggagc cctgtctcct gcctcccccg tcccaccact 420
 gaggtgtttg ggccaggcca gtcattgtag aacggcctcc tgcactggcc atgcccattgc 480
 ccacggaggg caccocgcca cctctgagt gcacccccat cccagtccca gcctacttcc 540
 gccacgcaga acctggattc tccctcaaga ggcccagggg gctcagccgg agcctccac 600
 ctccgcccc tgccaagggc agcattccca tcagccgcct cttccctcct cggaccccag 660
 gctggcacca gctgcagccc cggcgggtgt cattccgggg cgaggcctca gagactctgc 720
 agagccctgg gtatgaccca agccggccag agtccttctt ccagcagagc ttccagaggc 780
 tcagccgcct gggccatggc tccacaggag aggtcttcaa ggtgcgctcc aaggaggacg 840
 gccggtctta tgccgtaaaag cgttccatgt caccattccg gggccccaag gaccggggcc 900
 gcaagttggc cgaggtgggc agccacgaga aggtggggca gcacccatgc tgcgtgcggc 960
 tggagcaggc ctgggaggag ggcggcatcc tgtacctgca gacggagctg tgcggggcca 1020

gcctgcagca acactgtgag gcctgggggt ccagcctgcc tgaggccag gtctggggct 1080
 acctgcggga cacgtgtctt gccctggccc atctgcacag ccagggcctg gtgcaccttg 1140
 atgtcaagcc tgccaacatc ttcttggggc cccggggccg ctgcaagctg ggtgacttcg 1200
 gactgtgtgg ggagctgggt acagcaggag ctggtgaggt ccaggaggga gacccccgct 1260
 acatggcccc cgagctgctg cagggtctct atgggacagc agcggatgtg ttcagtctgg 1320
 gcctcaccat cctggaagtg gcatgcaaca tggagctgcc ccacgggtggg gagggtggc 1380
 agcagctgcg ccagggttac ctgccccctg agttcactgc cgtgtgtct tccagctgc 1440
 gttctgtcct tgtcatgatg ctggagccag accccaagct gcggggccac gccaggccc 1500
 tgctggcact gcctgtgttg aggcagccgc gggcctgggg tgtgctgtgg tgcatggcag 1560

```

cggaggccct gagccgaggg tgggccctgt ggcaggccct gcttgccctg ctctgctggc 1620
tctggcatgg gctggctcac cctgccagct ggctacagcc cctgggcccg ccagccaccc 1680
cgccctggctc accaccctgc agtttgctcc tggacagcag cctctccagc aactgggatg 1740
acgacagcct agggccttca ctctcccctg aggetgtcct ggcccggact gtggggagca 1800
cctccacccc ccggagcagg tgcacacca gggatgccct ggacctaatg gacatcaact 1860
cagagcctcc tcggggctcc tccccctcct ttgagcctcg gaacctctc agcctgtttg 1920
aggacaccct agacccaacc tgagccccag actctgcctc tgcactttta accttttatc 1980
ctgtgtctct cccgtcgccc ttgaaagctg gggcccctcg ggaactccca tggctctctc 2040
tgccctggccg tgtctaataa aaagtatttg aaccttggga gcaccaagc ttgctcatgt 2100
ggcaaaaaaa aaaaaaaaaa a 2121

```

<210> 138

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004203

<400> 138

ctggccgtgt ctaataaaaa gtatttgaac cttgggagca cccaagcttg ctcatgtggc 60

<210> 139

<211> 1982

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004207

<400> 139

```

ggcgagaggc gggtgagggc ggcccagcgg cggcagggtga ggcggaacca accctcctgg 60
ccatgggagg ggccgtggtg gacgagggcc ccacaggcgt caaggcccct gacggcggct 120
ggggctgggc cgtgctcttc ggctgtttcg tcatcactgg cttctcctac gccttcccca 180
aggccgtcag tgtcttcttc aaggagctca tacaggagtt tgggatcggc tacagcgaca 240
cagcctggat ctctccatc ctgctggcca tgctctacgg gacagggtccg ctctgcagtg 300
tgtgcgtgaa ccgctttggc tgccggcccg tcatgcttgt ggggggtctc tttgcgtcgc 360
tgggcatggg ggctgcgtcc ttttgccgga gcatcatcca ggtctacctc accactgggg 420
tcatcacggg gttgggtttg gcaactcaact tccagccctc gctcatcatg ctgaaccgct 480
acttcagcaa gcggcgcccc atggccaacg ggctggcggc agcaggtagc cctgtcttcc 540
tgtgtgccct gagcccgtg gggcagctgc tgcaggaccg ctacggctgg cggggcggct 600
tcctcctcct gggcgggcctg ctgctcaact gctgcgtgtg tgccgcactc atgaggcccc 660
tgggtgtcac ggcccagccg ggctcggggc cgcgcgacc ctcccggcgc ctgctagacc 720
tgagcgtctt ccgggaccgc ggctttgtgc ttacgccgt ggccgcctcg gtcattggtg 780
tggggctctt cgtcccgcgc gtgttcgtgg tgagctacgc caaggacctg ggcgtgcccg 840
acaccaaggc cgccttctg ctcaccatcc tgggcttcat tgacatcttc gcgcggccgg 900
ccgcgggctt cgtggcgggg cttgggaagg tgcggcccta ctccgtctac ctcttcagct 960
tctccatgtt cttcaacggc ctgcgggacc tggcgggctc tacggcgggc gactacggcg 1020
gcctcgtggg cttctgcac tcttttggca tctctacgg catggtgggg gccctgcagt 1080
tcgagggtgct catggccatc gtgggcaccc acaagttctc cagtgccatt ggcctgggtg 1140
tgctgatgga ggcggtggcc gtgctcgtcg ggcccccttc gggaggcaaa ctctgggatg 1200
cgacccacgt ctacatgtac gtgttcatcc tggcgggggc cgagggtgctc acctcctccc 1260
tgattttgct gctgggcaac ttcttctgca ttaggaagaa gcccaaagag ccacagcctg 1320
aggtggcggc cgcgaggagg gagaagctcc acaagcctcc tgcagactcg ggggtggact 1380
tgccgggagg ggagcatttc ctgaaggctg agcctgagaa aaacggggag gtggttcaca 1440
ccccggaaac aagtgtctga gtggctgggc ggggcggca ggcaagggg gagggtacag 1500
aagccggcaa cgcttctat ttattttaca aactggactg gctcaggcag ggccacggct 1560
gggctccagc tgccggccca gcggatcgtc gcccgatcag tgttttgagg ggggaagggtg 1620
cgggggtggg accgtgtcat tccagagtgg atctgcggtg aagccaagcc gcaaggttac 1680
aaggcatcct caccaggggc ccgcctgct gctcccagg gtgctgcggc cactgctatg 1740
ctcaaggacc tggaaaccca tgcttcgaga caacgtgact ttaatgggag ggtgggtggg 1800

```

```

ccgcagacag gctggcaggg caggtgctgc gtggggccct ctccagcccg tctaccctg 1860
ggctcacatg gggcctgtgc ccacccctct tgagtgtctt ggggacagct ctttccaccc 1920
ctggaagatg gaaataaacc tgcgtgtggg tggagtgttc tcgtgccgaa ttcaaaaagc 1980
tt 1982

```

```

<210> 140
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004207

```

```

<400> 140
cctcttgagt gtcttgggga cagctctttc cacccttgga agatggaaat aaacctgcgt 60

```

```

<210> 141
<211> 2054
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004209

```

```

<400> 141
cgggaggcgg cagcggctgc agcgttggtg gcatcagcat cagcatcagc ggcagcggca 60
gcggcctcgg gcggggcccg ccggacggac aggcggacag aaggcggcag gggcgcgcg 120
ccgccccggg ccggccatgg agggcgccct cttcggcgcg ggcggcgcgag gggccggcct 180
ggaccccggt agctttgcgc ggcgggccca gacctgctc cgggtcgcgt cctgggtggt 240
ctccatcgcc gtcttcgggc ccctcgtaaa cgagggctac gtgaacaccg acagcggccc 300
cgagctgcgc tgcgtgttca acgggaacgc gggcgccctc cgcttcggcg tcgcgttggg 360
cctcggagcc ttcctcgccg gcgcgcgctt cctgctgctc gatgtgcgct tccagcaaat 420
cagcagcgct cgcgaccgcc ggcgcgcggg gttgctggac ctgggcttct caggactctg 480
gtccttcctg tggttcgtgg gcttctgctt cctcaccat cagtggcagc gcacggcgcc 540
agggccggcc acgacgcagg cggggggacgc ggcgcggggc gccatcgctt cagcttctt 600
ctccatcctc agctgggtgg cgtcaccgt gaaggccctg cagcggttcc gcctgggcac 660
cgacatgtca ctcttcgcca ccgaacagct ggcacccggg gcgagccagg cctaccccg 720
ctatccggtg ggcagcggcg tggagggcac cgagacctac cagagccgc ccttcaccga 780
gacctggac accagccca aagggtacca ggtgcccgc tactagcggc tggcaggcac 840
agaccagggc tccaaggcca cccaccaaac gcaggcccca ggggtctcgg gacctccctt 900
gggtccttcc agctcagtc gcgcggacaga gtagggtggc gctttgogcc atccggggcc 960
aagagggggg ggaccgcgt gtctgggctg cccctgcca gttccccag tccctcagca 1020
cctggcccca ggactgaggt cctgagaagg ggatagcact gccaggagc tgtgtcccta 1080
gcctggaatg gactggcctg gggaaaggct tcccctcttg ggccacacct gctcactctg 1140
gggttggggg tccagctgcc ctctacgac aggtgcagg gctgcccagg acaaagcggg 1200
ggcaggggaa agacaccacc ctgcgcccaa gactggggat cctggccact gttcccatcc 1260
catgtccctg tgggtagtga ctgtctcgtt tctgtcatgg tgggtgcgtcc cgtccggagc 1320
cactctccac tttctctcac aggtgctag aacagcccag ccctgtcagt gttgtgatca 1380
tggtccagtc ttcgggtttc acctcctagt actccacaag ctgctcctct ctctgtggcc 1440
ccggccctg cccagggtg ggtgggtctg gccaggaagg cacaaggtag ctgtgggcca 1500
agacaccagc cctgtcctag cccttcagta agaccttgcc aggagaggag aaggatgcct 1560
gggtgccagg caagacaagc ccctcagcag gagagaggcc cagaggctcc agctggccac 1620
cgtgccccac aagatggccc ctgtgtggtt ccctttacct tggcttctct gccagtcctc 1680
tgctctcca cctgcacctt gcttctctgg ccagtcctag gttggagtc ctctgcatag 1740
ctgactactc atgcattgct caaagctggc ttttcacatt aagtcaacac caaacgtggt 1800
tgccacattt catcagacag acacctccct ctggagatgc agttgagtga caaccttggt 1860
acattgtagc ctagaccaat tctgtgtgga tatttaagt aacatgttta caatttttgt 1920
atatatcact ctctccctct cctgaaagac cagagattgt gtattttcag tgtcccatgt 1980
tccgactgca ccttctttac aataaagact gtaactgagc tgactgtgaa aaaaaaaaaa 2040
aaaaaaaaaa aaaa 2054

```

<210> 142
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004209

<400> 142
 gatgcagttg agtgacaacc ttgttacatt gtagcctaga ccaattctgt gtggatattt 60

<210> 143
 <211> 1224
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004217

<400> 143
 ggccgggaga gtagcagtgc cttggacccc agctctcctc cccctttctc tctaaggatg 60
 gccacagaagg agaactccta cccctggccc tacggccgac agacggctcc atctggcctg 120
 agcaccctgc cccagcgagt cctccggaaa gagcctgtca ccccatctgc acttgtcctc 180
 atgagccgct ccaatgtcca gccacacagct gcccctggcc agaaggtgat ggagaatagc 240
 agtgggacac ccgacatctt aacgcggcac ttcacaattg atgactttga gattgggcgt 300
 cctctgggca aaggcaagtt tggaaacgtg tacttggctc gggagaagaa aagccatttc 360
 atcgtggcgc tcaaggtcct cttcaagtcc cagatagaga aggagggcgt ggagcatcag 420
 ctgcgagag agatcgaaat ccaggccccac ctgcaccatc ccaacatcct gcgtctctac 480
 aactatTTTT atgaccggag gaggatctac ttgattctag agtatgcccc ccgcggggag 540
 ctctacaagg agctgcagaa gagctgcaca tttgacgagc agcgaacagc cagcatcatg 600
 gaggagttgg cagatgctct aatgtactgc catgggaaga aggtgattca cagagacata 660
 aagccagaaa atctgtctct agggctcaag ggagagctga agattgctga cttcggctgg 720
 tctgtgcatg cgccctccct gaggaggaag acaatgtgtg gcaccctgga ctacctgcc 780
 ccagagatga ttgaggggcg catgcacaat gagaaggtgg atctgtgttg cattggagtg 840
 ctttgctatg agctgctggg ggggaaccca ccttttgaga gtgcatacaca caacgagacc 900
 tatcgccgca tcgtcaagggt ggacctaaag ttcccgcctt ctgtgccac gggagcccag 960
 gacctcatct ccaaactgct caggcataac ccctcggaac ggctgcccc ggcccagggtc 1020
 tcagcccacc cttgggtccg ggccaactct cggaggggtgc tgccctccctc tgcccttcaa 1080
 tctgtgcct gatgggtccct gtcattcact cgggtgcgtg tgtttgtatg tctgtgtatg 1140
 tataggggaa agaagggatc cctaactgtt cccttatctg ttttctacct cctcctttgt 1200
 ttaataaagg ctgaagcttt ttgt 1224

<210> 144
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004217

<400> 144
 gtctgtgtat gtatagggga aagaaggat ccctaactgt tcccttatct gttttctacc 60

<210> 145
 <211> 983
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004335

<400> 145
 gtggaattca tggcatctac ttctgtatgac tattgcagag tgcccatgga agacggggat 60
 aagcgtgtga agctttctgct ggggatagga attctgggtgc tcctgatcat cgtgattctg 120
 ggggtgccct tgattatctt caccatcaag gccaacagcg aggcctgccg ggacggcctt 180
 cgggcagtga tggagtgtcg caatgtcacc catctcctgc aacaagagct gaccgaggcc 240
 cagaagggtt ttcaggatgt ggaggcccag gccgccacct gcaaccacac tgtgatggcc 300
 ctaatggctt ccctggatgc agagaaggcc caaggacaaa agaaagtggg ggagcttgag 360
 ggagagatca ctacattaaa ccataagctt caggacgcgt ctgcagagggt ggagcgactg 420
 agaagagaaa accagggtctt aagcgtgaga atcgccggaca agaagtacta cccagctcc 480
 caggactcca gctccgctgc ggcgcccag ctgctgattg tgctgctggg cctcagcgct 540
 ctgctgcagt gagatcccag gaagctggca catcttggaa ggtccgtcct gctcggcttt 600
 tcgcttgaac attcccttga tctcatcagt tctgagcggg tcatggggca acacgggttag 660
 cggggagagc acggggtagc cggagaaggg cctctggagc aggtctggag gggccatggg 720
 gcagtcctgg gtgtggggac acagtcgggt tgacccaggg ctgtctccct ccagagcctc 780
 cctccggaca atgagtcccc cctcttgtct cccaccctga gattgggcat ggggtgcggg 840
 gtggggggcga tgtgctgcct gttgttatgg gttttttttg cggggggggg tgcttttttc 900
 tgggggtcttt gagctccaaa aaataaacac ttcttttgag ggagagcaaa aaaaaaaaaa 960
 aaaaaaaaaa aaaaaaaaaa aaa 983

<210> 146
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004335

<400> 146
 ggttgctttt ttctgggggtc tttgagctcc aaaaaataaa cacttccttt gagggagagc 60

<210> 147
 <211> 3446
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004336

<400> 147
 ttctagtttt cggttcaggt ttgccgctgc cggccagcgt cctctggcca tggacacccc 60
 ggaaaatgtc cttcagatgc ttgaagccca catgcagagc tacaagggca atgaccctct 120
 tgggtgaatgg gaaagataca tacagtgggt agaagagaat tttcctgaga ataaagaata 180
 cttgataact ttactagaac atttaaatgaa ggaattttta gataagaaga aataccacaa 240
 tgacccaaga ttcacagttt attgttttaa atttgctgag tacaacagtg acctccatca 300
 atttttttgag tttctgtaca accatgggat tggaaacctg tcatcccctc tgtacattgc 360
 ctgggcgggg catctggaag cccaaggaga gctgcagcat gccagtgcctg tccttcagag 420
 aggaattcaa aaccaggctg aaccagaga gttcctgcaa caacaataca ggttatttca 480
 gacacgcctc actgaaaccc atttgccagc tcaagctaga acctcagaac ctctgcataa 540
 tgttcagggt tttaatcaaa tgataacatc aaaatcaaat ccaggaaata acatggcctg 600
 cattttctaag aatcagggtt cagagctttc tggagtata tcttcagctt gtgataaaga 660
 gtcaaatatg gaacgaagag tgatcacgat ttctaaatca gaatattctg tgcactcatc 720
 tttggcatcc aaagttgatg ttgagcaggt tggtatgtat tgcaaggaga agcttattcg 780
 tggggaatca gaattttcct ttgaagaatt gagagcccag aaatacaatc aacggagaaa 840
 gcatgagcaa tgggtaaatg aagacagaca ttatatgaaa aggaagaag caaatgcttt 900
 tgaagaacag ctattaaaac agaaaatgga tgaacttcac aagaagttgc atcaggtggg 960
 ggagacatcc catgaggatc tgcccgcttc ccaggaaagg tccgagggtta atccagcacg 1020
 tatggggcca agtgtaggct cccagcagga actgagagcg ccatgtcttc cagtaacctc 1080
 tcagcagaca ccagtgaaca tggaaaagaa cccaagagag gcacctcctg ttgttcctcc 1140
 tttggcaaat gctattttctg cagcttttgt gtccccagcc accagccaga gcattgctcc 1200
 tcctgttcct ttgaaagccc agacagtaac agactccatg tttgcagtgg ccagcaaaga 1260
 tgctggatgt gtgaataaga gtactcatga attcaagcca cagagtggag cagagatcaa 1320

```

agaaggggtgt gaaacacata aggttgccaa cacaagttct tttcacacaa ctccaaacac 1380
atcactggga atggttcagg caacgccatc caaagtgcag ccatcaccca cegtgcacac 1440
aaaagaagca ttaggtttca tcatgaatat gtttcaggct cctacacttc ctgatatttc 1500
tgatgacaaa gatgaatggc aatctctaga tcaaaatgaa gatgcatttg aagcccagtt 1560
tcaaaaaaat gtaaggatcat ctggggccttg gggagtcatt aagatcatct cttctttgtc 1620
atctgctttt catgtgtttg aagatggaaa caaagaaaat tatggattac cacagcctaa 1680
aaataaaccc acaggagcca ggaccttttg agaacgctct gtcagcagac ttccttcaaa 1740
accaaaaggag gaagtgcctc atgctgaaga gtttttggat gactcaactg tatggggtat 1800
tcgctgcaac aaaaccctgg caccagtcct taagagccca ggagacttca catctgctgc 1860
acaacttgcg tctacacctt tccacaagct tcagtgagg tccagtgcac ttttagaaga 1920
taaagaaaaat gtggtagcaa aacagtgtac ccaggcgact ttggattctt gtgaggaaaa 1980
catggtgggtg ccttcaaggg atggaaaatt cagtccaatt caagagaaaa gcccaaaaca 2040
ggccttgtcg tctcacatgt attcagcatc cttacttctg ctgagccagc ctgctgcagg 2100
tggggctactt acctgtgagg cagagttggg cgttgaggct tgcagactca cagacactga 2160
cgctgccatt gcagaagatc caccagatgc tattgctggg ctccaagcag aatggatgca 2220
gatgagttca cttgggactg ttgatgctcc aaacttcatt gttgggaacc catgggatga 2280
taagctgatt ttcaaacttt tatctgggct ttctaaacca gtgagttcct atccaaatac 2340
ttttgaatgg caatgtaaac ttccagccat caagcccaag actgaatttc aattgggttc 2400
taagctggtc tatgtccatc accttcttgg agaaggagcc tttgccaggg tgtacgaagc 2460
taccaggga gatctgaatg atgctaaaaa taaacagaaa tttgttttaa aggtccaaaa 2520
gcctgccaac cctgggaat tctacatttg gaccagttg atggaaagac taaagccatc 2580
tatgcagcac atgtttatga agttctattc tgccactta ttccagaatg gcagtgtatt 2640
agtaggagag ctctacagct atggaacatt attaaatgcc attaacctct ataaaaatac 2700
ccctgaaaaa gtgatgcctc aaggctctgt catctctttt gctatgagaa tgctttacat 2760
gattgagcaa gtgcatgact gtgaaatcat tcatggagac attaaaccag acaatttcat 2820
acttgaaaac ggatttttgg aacaggatga tgaagatgat ttatctgctg gcttggcact 2880
gattgacctg ggtcagagta tagatatgaa actttttcca aaaggaacta tattcacagc 2940
aaagtgtgaa acatctggtt ttcatgtgtg tgagatgctc agcaacaaac catggaacta 3000
ccagatcgat tactttgggg ttgctgcaac agtatattgc atgctctttg gcactttacat 3060
gaaagtgaaa aatgaaggag gagagtgtaa gcctgaaggc ctttttagaa ggcttctca 3120
tttgatgatg tggaatgaat tttttcatgt tatgttgaat attccagatt gtcacatct 3180
tccatctttg gatgtgttaa ggcaaaagct gaagaaagta tttcaacaac actatactaa 3240
caagattagg gccctacgta ataggcta atgtactgctc ttagaatgta agcggttcaag 3300
aaaataaaat ttggatatag acagtcttta aaaatcacac tgtaaatatg aatctgctca 3360
ctttaaacct gttttttttt catttattgt ttatgtaaat gtttgttaaa aataaatccc 3420
atggaatatt tccatgtaaa aaaaaa 3446

```

<210> 148
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004336

<400> 148
 ttagggccct acgtaatagg ctaattgtac tgctcttaga atgtaagcgt tcacgaaaat 60

<210> 149
 <211> 739
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004345

<400> 149
 taaagcaaac ccagccac accctggcag gcagccagg atgggtggat caggaaggct 60
 cctgggtggg cttttgcatc aggtcaggc tgggcataaa ggaggctcct gtgggctaga 120
 gggaggcaga catggggacc atgaagacc aaagggatgg ccactccctg gggcggtggt 180
 cactggtgct cctgctgctg ggctggtga tgctctggc catcattgcc caggtcctca 240

```

gctacaagga agctgtgctt cgtgctatag atggcatcaa ccagcgggtcc tcggatgcta 300
acctctaccg cctcctggac ctggacccca ggcccacgat ggatggggac ccagacacgc 360
caaagcctgt gagcttcaca gtgaaggaga cagtgtgccc caggacgaca cagcagtcac 420
cagaggattg tgacttcaag aaggacgggc tggatgaagc gtgtatgggg acagtgacct 480
tcaaccaggc caggggctcc tttgacatca gttgtgataa ggataacaag agatttgccc 540
tgctgggtga tttcttccgg aaatctaaag agaagattgg caaagagttt aaaagaattg 600
tccagagaat caaggatttt ttgcggaatc ttgtaccag gacagagtcc tagtgtgtgc 660
cctaccctgg ctacggcttc tgggctctga gaaataaact atgagagcaa tttcaaaaaa 720
aaaaaaaaa aaaaaaaaaa 739

```

<210> 150
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004345

```

<400> 150
gcaaagagtt taaaagaatt gtccagagaa tcaaggattt tttgcggaat cttgtaccca 60

```

<210> 151
 <211> 1432
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004577

```

<400> 151
gaggaaaatt cttccagcga tgggtctccca ctacagagctg aggaagcttt tctactcagc 60
agatgctgtg tgttttgatg ttgacagcac ggtcatcaga gaagaaggaa tcgatgagct 120
agccaaaatc tgtggcggtt aggacgcggt gtcagaaatg acacggcgag ccatgggcgg 180
ggcagtgcc ttcaaaagctg ctctcacaga gcgcttagcc ctcatccagc cctccaggga 240
gcagggtgcag agactcatag cagagcaacc cccacacctg acccccgga taaggagct 300
ggtaagtgcg ctacaggagc gaaatgttca ggttttccta atatctggtg gctttaggag 360
tattgtagag catgttgctt caaagctcaa tatccagca accaatgtat ttgccaatag 420
gctgaaattc tactttaacg gtgaatatgc aggttttgat gagacgcagc caacagctga 480
atctgggtgga aaaggaaaag tgattaaact tttaaaggaa aaatttcatt ttaagaaaat 540
aatcatgatt ggagatgggtg ccacagatat ggaagcctgt cctcctgctg atgctttcat 600
tggatttgga ggaaatgtga tcaggcaaca agtcaaggat aacgccaaat ggtatatcac 660
tgattttgta gagctgctgg gagaactgga agaataacat ccattgtcgt acagctcaa 720
acaacttcag atgaattttt acaagttata cagattgata ctgtttgctt acagttgcct 780
attacaactt gctatagaaa gttggtacaa atgatctgta ctttaacta cagttaggaa 840
tcctagaaga ttgctttttt ttttttttta actgtagtcc cagtattata tgatgactat 900
tgatttcctg gagaggtttt tttttttttt gagacagaa cttgtctgtg tgcccaggct 960
ggagtgcagt ggcgcggtct cggtcactg caagctctgc ctcccagggt cagccattc 1020
tcctgcctca gcctcccgag tagctgggac tacaggcacc cgccaccaca tccggcta 1080
tttttgatt tttagtagag acgggggttg accgtgtag ccaggatggt cttgatctcc 1140
tgacctgtg atccgcctgc ctacagctcc caaagtgtg ggattacagg cttgggccac 1200
cgccccagc caatgtccta gagagttttg tgatctgaat tctttatgta tttttgtagc 1260
tatatttcat acaaagtgtc ttaagtgtgg agagtcaatt aaacacctt actottagaa 1320
atacggattc ggcagccttc agtgaatatt ggtttctctt tggtagtca ataaaagttt 1380
atccgtatgt cagaacggat ttgtggaaaa aaaaaaaaaa aaaaaaaaaa aa 1432

```

<210> 152
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_004577

<400> 152

tagaaatcacg gattcggcag ccttcagtga atattggttt ctctttggta tgtcaataaa 60

<210> 153

<211> 1530

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004701

<400> 153

```

aatcctggaa caaggctaca gcgtcgaaga tccccagcgc tgcgggctcg gagagcagtc 60
ctaacggcgc ctcgtacgct agtgtcctcc cttttcagtc cgcgtccctc cctgggcccgg 120
gctggcactc ttgccttccc cgteccctcat ggcgctgctc cgacgcccga cgggtgtccag 180
tgattttggag aatattgaca caggagttaa ttctaaagtt aagagtcattg tgactattag 240
gcgaactggt ttagaagaaa ttggaaatag agttacaacc agagcagcac aagtagctaa 300
gaaagctcag aacaccaaag ttccagttca accaccaa acaacaaatg tcaacaaaca 360
actgaaacct actgcttctg tcaaaccagt acagatggaa aagttggctc caaaggggtcc 420
ttctcccaca cctgaggatg tctccatgaa ggaagagaat ctctgccaaag ctttttctga 480
tgctttgctc tgcaaaatcg aggacattga taacgaagat tgggagaacc ctcagctctg 540
cagtactac gttaaggata tctatcagta tctcaggcag ctggaggttt tgcagtcctat 600
aaacccacat ttcttagatg gaagagatat aaatggacgc atgcgtgccca tcttagtgga 660
ttggctggta caagtccact ccaagtttag gcttctgcag gagactctgt acatgtgctg 720
tggtattatg gatcgatttt tacaggttca gccagtttcc cggaagaagc ttcaattagt 780
tggtgattact gctctgctct tggcttccaa gtatgaggag atgttttctc caaatattga 840
agactttgtt tacatcacag acaatgctta taccagttcc caaatccgag aaatggaaac 900
tctaattttg aaagaattga aatttgagtt gggctcgacc ttgccactac acttcttaag 960
gcgagcatca aaagccgggg aggttgatgt tgaacagcac actttagcca agtatttgat 1020
ggagctgact ctcatcgact atgatatggt gcattatcat ccttctaagg tagcagcagc 1080
tgcttctctg ttgtctcaga aggttctagg acaaggaaaa tggaaactta agcagcagta 1140
ttacacagga tacacagaga atgaagtatt ggaagtcatt cagcacatgg ccaagaatgt 1200
ggtgaaagta aatgaaaact taactaaatt catcgccatc aagaataagt atgcaagcag 1260
caaactcctg aagatcagca tgatccctca gctgaactca aaagccgtca aagaccttgc 1320
ctcccactg ataggaaggt cctaggctgc cgtgggcccgt ggggatgtgt gcttcattgt 1380
gccctttttc ttattggttt agaactcttg attttgtaca tagtctctctg gtctatctca 1440
tgaaacctct tctcagacca gttttctaaa catatattga ggaaaaataa agcgattgggt 1500
ttttcttaag gtaaaaaaaaa aaaaaaaaaa 1530

```

<210> 154

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004701

<400> 154

agaactcttg attttgtaca tagtctctctg gtctatctca tgaaacctct tctcagacca 60

<210> 155

<211> 2536

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004702

<400> 155


```

agcgggtgcg gggcgggacc ggccccggcct atatatattggg ttggcgccgg cgccagctga 60
gccgagcggt agctgggtctg gcgaggtttt atacacctga aagaagagaa tgtcaagacg 120
aagtagccgt ttacaagcta agcagcagcc ccagcccagc cagacggaat cccccaaga 180
agcccagata atccaggcca agaagaggaa aactaccag gatgtcaaaa gaagtctggc 240
taaakatgtt aaaaaaggag agcagatatg ttcatgacaa acattttgaa gttctgcatt 300
ctgacttggg accacagatg aggtccatac ttctagactg gcttttagag gtatgtgaag 360
tatacacact tcatagggaa acattttatc ttgcacaaga cttttttgat agatttatgt 420
tgacacaaaa ggatataaat aaaaatatgc ttcaactcat tgggaattacc tcattattca 480
ttgcttccaa acttgaggaa atctatgctc ctaaactcca agagtttgct tacgtcactg 540
atggtgcttg cagtgaagag gatattctaa ggatggaaact cattatatta aaggctttta 600
aatgggaact ttgtcctgta acaatcatct cctggctaaa tctctttctc caagttgatg 660
ctcttaaaaga tgctcctaaa gttctttctac ctcatgtatc tcaggaaaca ttcatcctaa 720
tagctcagct tttagatctg tgtattctag ccattgattc attagagttc cagtacagaa 780
tactgactgc tgctgccttg tgccatttta cctccattga agtgggttaag aaagcctcag 840
gtttggagtg ggacagtatt tcagaatgtg tagattggat ggtacccttt gtcaatgtag 900
taaaaagtac tagtccagtg aagctgaaga cttttaagaa gattcctatg gaagacagac 960
ataatatcca gacacatata aactatttgg ctatgctgga ggaagtaaat tacataaaca 1020
ccttcagaaa aggggggacag ttgtcaccag tgtgcaatgg aggcattatg acaccaccga 1080
agagcactga aaaaccacca ggaaaacact aaagaagata actaagcaaa caagttggaa 1140
ttcaccaaga ttgggtagaa ctggtatcac tgaactacta aagttttaca gaaagtagtg 1200
ctgtgattga ttgccttagc caattcacaa gttacactgc cattctgatt ttaaaactta 1260
caattggcac taaagaatac atttaattat ttctatggt agctgtttaa gaaacagcag 1320
gacttgttta caaagatgtc ttcatccca aggttactgg atagaagcca accacagtct 1380
ataccatagc aatgtttttc cttaaatcca gtgttactgt gtttatcttg ataaactagg 1440
aattttgtca ctggagtttt ggactggata agtgctacct taaaggggat actaagtgat 1500
acagtacttt gaatctagtt gttagattct caaaattcct acactcttga ctagtgcaat 1560
ttggttcttg aaaattaaat ttaaacttgt ttacaaaggt ttagttttgt aataaggtga 1620
ctaatttatc tatagctgct atagcaagct attataaaac ttgaatttct acaaaggtg 1680
aaatttaatg ttttttaaac tagtttattt gccttgccat aacacatttt ttaactaata 1740
aggcttagat gaacatggg ttcaacctgt gctctaaaca gtgggagtag caaagaaatt 1800
ataaacaaga taaatgctgt ggctccttcc taactggggc tttcttgaca tgtaggttg 1860
ttggtaataa ccttttttga tatcacaatt tgggtgaaaa acttaagtac cctttcaaac 1920
tatttatatg aggaagtcac tttactactc taagatatcc ctaaggaatt tttttttta 1980
athtagtggt actaaggctt tatttatggt tgtgaaactg ttaaggtcct ttctaaattc 2040
ctccattgtg agataaggac agtgtcaaag tgataaagct taacacttga cctaaacttc 2100
tattttctta aggaagaaga gtattaaata tatactgact cctagaaatc tatttattaa 2160
aaaaagacat gaaaacttgc tgtacatagg ctactatatt ctaaataatt taaattagct 2220
tttctaaaaa aaaaatccag cctcataaag tagattagaa aactagattg ctagttttatt 2280
ttgttatcag atatgtgaat ctcttctccc tttgaagaaa ctatacattt attgttacgg 2340
tatgaagtct tctgtatagt ttgtttttta actaatattt gtttcagtat tttgtctgaa 2400
aagaaaacac cactaattgt gtacatatgt attatataaa cttaaccttt taatactgtt 2460
tatttttagc ccattgttta aaaaataaaa gttaaaaaaa tttaactgct taaaagtaaa 2520
aaaaaaaaa aaaaaa 2536

```

<210> 156

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004702

<400> 156

gtttgtgaaa ctgttaaggc cttttctaaa ttctccatt gtgagataag gacagtgta 60

<210> 157

<211> 1491

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004710

<400> 157

```

gcgggcgggcg cagcgggcggc gacgggcgaca tggagagcgcg ggcctacggc gcggccaagg 60
cgggcgggctc cttcgacctg cggcgcttcc tgacgcagcc gcagggtggcg gcgcgcgcgc 120
tgtgcttgggt cttcgcccttg atcgtgttct cctgcactca tgggtgagggc tacagcaatg 180
cccacgagtc taagcagatg tactgcgtgt tcaaccgcaa cgaggatgcc tgccgctatg 240
gcagtgccat cgggggtgctg gccttcctgg cctcggcctt cttcttggcg gtcgacgcgt 300
atctcccccga gatcagcaac gccactgacc gcaagtacct ggtcattggg gacctgctct 360
tctcagctct ctggaccttc ctgtggtttg ttggtttctg cttcctcacc aaccagtggg 420
cagtcaccaa ccgaaggac gtgctgggtg gggccgactc tgtgagggca gccatcacct 480
tcagcttctt ttccatcttc tctgggggtg tgctggcctc cctggcctac cagcgctaca 540
aggtctggcg ggacgacttc atccagaatt acgttgaccc cactccggac cccaacactg 600
cctacgcctc ctaccaggt gcactctgtg acaactacca acagccacc ttccccaga 660
acgcggagac caccgagggc taccagccgc cccctgtgta ctgagcggcg gttagcgtgg 720
gaaggggggac agagaggggc ctccccctctg ccttggaact tcccatgagc ctctggaac 780
tgccagcccc tctctttcac ctgttccatc ctgtgcagct gacacacagc taaggagcct 840
catagcctgg cgggggctgg cagagccaca cccaagtg cgtgtgccag agggcttcag 900
tcagccgctc actcctccag ggcattttta ggaaaggggt ttcagctagt gtttttctc 960
gcttttaatg acctcagccc cgcctgcagt ggctagaagc cagcaggtgc ccatgtgcta 1020
ctgacaagtg cctcagcttc cccccggccc gggtcaggcc gtgggagccg ctattatctg 1080
cgttctctgc caaagactcg tgggggccc atcacctgcc ctgtgcagcg gagccggacc 1140
aggtctctgt gtcctcactc aggtttgctt cccctgtgcc cactgctgta tgatctgggg 1200
gccaccacc cgtgcgggtg gcctctgggc tgccctccgt ggtgtgaggg cggggctggg 1260
gctcatggca cttcctcctt gctccacccc ctggcagcag ggaagggctt tgcctgacaa 1320
caccagctt tatgtaaata ttctgcagtt gttacttagg aagcctgggg agggcagggg 1380
tgccccatgg ctcccagact ctgtctgtgc cgagtgtatt ataaaatcgt gggggagatg 1440
ccgggcctgg gatgtgttt ggagacggaa taaatgttt ctcattcagt a 1491

```

<210> 158

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004710

<400> 158

```

ttgctgaca acaccagct ttatgtaaatt attctgcagt tggtacttag gaagcctggg 60

```

<210> 159

<211> 3324

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004856

<400> 159

```

gcagagcacc gcgccttagc cgcgaagttc tagttcttgc tgccggctct aacgtccgc 60
agtcttcgcc agccagccgt cccgcagtcg cgtttgggcg gcgtggagcc tgctgccatg 120
aagtcagcga gagctaagac accccgga cctaccgtga aaaaagggtc ccaaaccgaa 180
cttaaagacc cagttgggggt aactgttagg gtgcgcccac tgggctttcc tgatcaagag 240
tggtgcatag aagtgatcaa taatacaact gttcagcttc atactcctga gggctacaga 300
ctcaaccgaa atggagacta taaggagact cagtattcat ttaaacaagt atttggcact 360
cacaccacc agaaggaaact ctttgatgtt gtggctaata ccttgggtcaa tgacctcatt 420
catggcaaaa atggctcttct ttttacatat ggtgtgacgg gaagtggaaa aactcacaca 480
atgactgggt ctccagggga agggggctg cttcctcgtt gtttggacat gatctttaac 540
agtatagggt catttcaagc taaacgatat gttttcaa ataatgatag gaatagtag 600
gatatacagt gtgaggttga tgccttatta gaacgtcaga aaagagaagc tatgcccaat 660
ccaaagactt cttctagcaa acgacaagta gatccagagt ttgcagatat gataactgta 720

```

```

caagaattct gcaaagcaga agaggttgat gaagatagtg tctatgggtg atttgtctct 780
tatattgaaa tatataataa ttacatatat gatctattgg aagaggtgcc gtttgatccc 840
ataaaaccca aacctccaca atctaaattg cttcgtgaag ataagaacca taacatgtat 900
gttgcaggat gtacagaagt tgaagtgaaa tctactgagg aggcttttga agttttctgg 960
agaggccaga aaaagagagc tattgctaatt acccatttga atcgtgagtc cagccgttcc 1020
catagcgtgt tcaacattaa attagttcag gctcccttgg atgcagatgg agacaatgtc 1080
ttacaggaaa aagaacaaat cactataagt cagttgtcct tggtagatct tgctggaagt 1140
gaaagaacta accggaccag agcagaaggg aacagattac gtgaagctgg taatattaat 1200
cagtcactaa tgaogctaag aacatgtatg gatgtcctaa gagagaacca aatgtatgga 1260
actaacaaga tgggtccata tgcagattca aagttaacc atctgttcaa gaactacttt 1320
gatggggaag gaaaagtgcg gatgatcgtg tgtgtgaacc ccaaggctga agattatgaa 1380
gaaaacttgc aagtcattgag atttgcggaa gtgactcaag aagttgaagt agcaagacct 1440
gtagacaagg caatatgtgg tttaacgcct gggaggagat acagaaacca gcctcgaggt 1500
ccagttggaa atgaaccatt ggttactgac gtgggttttgc agagttttcc acctttgccg 1560
tcatgcgaaa ttttggtatc caacgatgag cagacacttc caaggctgat tgaagcctta 1620
gagaaacgac ataacttacg acaaatgatg attgatgagt ttaacaaaca atctaattgt 1680
tttaaagctt tgttacaaga atttgacaat gctgttttaa gtaaagaaaa ccacatgcaa 1740
gggaaaactaa atgaaaagga gaagatgac tccaggacaga aattggaaat agaacgactg 1800
gaaaagaaaa acaaaacttt agaataatag attgagattt tagagaaaac aactactatc 1860
tatgaggaag ataaacgcaa tttgcaacag gaacttgaaa ctcagaacca gaaacttcag 1920
cgacagtttt ctgacaaacg cagattagaa gccaggttgc aaggcatggg gacagaaacg 1980
acaatgaagt gggagaaaaga atgtgagcgt agagtggcag ccaaacagct ggagatgcag 2040
aataaactct ggggttaaaga tgaaaagctg aaacaactga aggtctattgt tactgaacct 2100
aaaactgaga agccagagag acctctcgg gagcgagatc gagaaaaagt tactcaaaga 2160
tctgtttctc catcacctgt gcccttactc tttcaacctg atcagaacgc accaccaatt 2220
cgtctccgac acagacgac acgctctgca ggagacagat gggtagatca taagcccgcc 2280
tctaactatc aaactgaaac agtcattgcag ccacatgtcc ctcatgccat cacagtatct 2340
gttgcaaatg aaaaggcact agctaagtgt gagaagtaca tgctgaccca ccaggaacta 2400
gcctccgatg gggagattga aactaaacta attaaggggt atatttataa aacaaggggt 2460
ggtggacaat ctgttcagtt tactgatatt gagactttta agcaagaatc accaaatgg 2520
agtcgaaaaac gaagatcttc cacagtagca cctgcccac cagatgggtgc agagtctgaa 2580
tggaaccgat tagaaacaag gtgttctgtg gctgtggaga tgagagcagg atcccagctg 2640
ggacctggat atcagcatca cgcacaaccc aagcgcaaaa agccatgaac tgacagtccc 2700
agtactgaaa gaacattttc atttgtgtgg atgattttct gaaagccatg ccagaagcag 2760
tcttccaggt catctttagt aactccagct ttgttgaaaa tcacggacct cagctacatc 2820
atacactgac ccagagcaaa gctttcccta tggttccaaa gacaactagt attcaacaaa 2880
ccttgatatag tatatgtttt gccatattta atattaatag cagaggaaga ctcttttttt 2940
catcactgta tgaatttttt ataattgttt tttaaaatat atttcatgta tacttataaa 3000
ctaattcaca caagtgtttg tcttagatga ttaagggaaga ctatatctag atcatgtctg 3060
attttttatt gtgacttctc cagccctggg ctgaattttc taagggtttt taaacaaatg 3120
ctgctattta ttagctgcaa gaatgcactt tagaactatt tgacaattca gactttcaaa 3180
ataaagatgt aaatgactgg ccaataataa ccattttagg aaggtgtttt gaattctgta 3240
tgtatatatt cactttctga cattttagata tgccaaaaga attaaaatca aaagcactaa 3300
gaaataaaaa aaaaaaaaaa aaaa 3324

```

<210> 160

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004856

<400> 160

caaagctttc cctatggttc aaagacaact agtattcaac aaaccttgta tagtgtatgt 60

<210> 161

<211> 1536

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004900

<400> 161

```

acagagcttc aaaaaaagag cgggacaggg acaagcgtat ctaagaggct gaacatgaat 60
ccacagatca gaaatccgat ggagcggatg tatcgagaca cattctacga caactttgaa 120
aacgaaccca tcctctatgg tcggagctac acttggctgt gctatgaagt gaaaataaag 180
agggggccgt caaatctcct ttgggacaca ggggtcttct gaggccaggt gtatttcaag 240
cctcagtacc acgcagaaat gtgcttcctc tcttggttct gtggcaacca gctgcctgct 300
tacaagtgtt tccagatcac ctggtttgta tcttggaacc cctgcccga ctgtgtggcg 360
aagctggccg aattcctgtc tgagcacccc aatgtcacc tgaccatctc tgccgccgc 420
ctctactact actgggaaag agattaccga agggcgctct gcaggctgag tcaggcagga 480
gcccgcgtga cgatcatgga ctatgaagaa tttgcatact gctgggaaaa ctttgtgtac 540
aatgaaggct agcaattcat gccttggtac aaattcgatg aaaattatgc attcctgcac 600
cgcacgctaa aggagattct cagatacctg atggatccag acacattcac tttcaacttt 660
aataatgacc ctttggctct tcgacggcgc cagacctact tgtgctatga ggtggagcgc 720
ctggacaatg gcacctgggt cctgatggac cagcacatgg gctttctatg caacgaggct 780
aagaatcttc tctgtggctt ttacggccgc catgcgagc tgcgcttctt ggacctggtt 840
ccttcttttg agttggaccc ggcccagatc tacagggtca cttgggtcat ctctgggagc 900
ccctgcttct cctggggctg tgccggggaa gtgcgtgctg tccttcagga gaacacacac 960
gtgagactgc gcatcttcgc tgcccgcac tatgattacg accccctata taaggaggcg 1020
ctgcaaatgc tgccgggatg tggggcccaa gtctccatca tgacctacga tgagtttgag 1080
tactgctggg acacctttgt gtaccgccag ggatgtccct tccagccctg ggatggacta 1140
gaggagcaca gccaaagccct gagtgggagg ctgcccggca ttctccagaa tcagggaac 1200
tgaaggatgg gcctcagtct ctaaggaagg cagagacctg ggttgagcag cagaataaaa 1260
gatcttcttc caagaaatgc aaacagaccc ttcaccacca tctccagctg ctacagaca 1320
ccagcaaagc aatgtgctcc tgatcaagta gattttttaa aaatcagagt caattaattt 1380
taattgaaaa tttctcttat gttccaagtg tacaagagta agattatgct caatattccc 1440
agaatagttt tcaatgtatt aatgaagtga ttaattggct ccatatttag actaataaaa 1500
cattaagaat cttccataat tgtttccaca aacact 1536

```

<210> 162

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004900

<400> 162

```

tgctcacaga caccagcaaa gcaatgtgct cctgatcaag tagatttttt aaaaatcaga 60

```

<210> 163

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004988

<400> 163

```

cgtagagttc ggccgaagga acctgaccca ggctctgtga ggaggcaagg ttttcagggg 60
acaggccaac ccagaggaca ggattccctg gaggccacag aggagcacca aggagaagat 120
ctgcctgtgg gtcttcattg cccagctcct gccacactc ctgcctgctg ccctgacgag 180
agtcatcatg tctcttgagc agaggagtct gcactgcaag cctgaggaag cccttgaggc 240
ccaacaagag gccctgggcc tgggtgtgtg gcaggctgcc gcctcctcct cctctcctct 300
ggctcctggg accctggagg aggtgcccac tgctgggtca acagatcctc cccagagtcc 360
tcaggagacc tccgcctttc ccactaccat caacttcact cgacagaggc aaccagtgga 420
gggttccagc agccgtgaag aggaggggcc aagcacctct tgtatcctgg agtccttggt 480
ccgagcagta atcactaaga aggtggctga tttggttggg tttctgctcc tcaaataatcg 540
agccagggag ccagtcacaa aggcagaaat gctggagagt gtcacaaaa attacaagca 600

```

```

ctgttttctt gagatcttcg gcaaagcctc tgagtccctg cagctgggtct ttggcattga 660
cgtgaaggaa gcagacccca ccggccactc ctatgtcctt gtcacctgcc taggtctctc 720
ctatgatggc ctgctgggtg ataatcagat catgcccaag acaggcttcc tgataattgt 780
cctgggtcatg attgcaatgg agggcggcca tgctcctgag gaggaaatct gggaggagct 840
gagtgtgatg gaggtgtatg atgggagggg gcacagtgcc tatggggagc ccaggaagct 900
gctcaccxaa gatttggtgc aggaaaagta cctggagtac cggcaggtgc cggacagtga 960
tcccgcacgc tatgagttcc tgtgggggtc aagggccctt gctgaaacca gctatgtgaa 1020
agtccttgag tatgtgatca aggtcagtgc aagagtctgc tttttcttcc catccctgcg 1080
tgaagcagct ttgagagagg aggaagaggg agtctgagca tgagttgcag ccaggggccag 1140
tgaggagggg actggggccag tgcaccttcc agggccgcgt ccagcagctt ccctgcctc 1200
gtgtgacatg agggccattc ttcactctga agagagcggg cagtgttctc agtagtaggt 1260
ttctgttcta ttgggtgact tggagattta tctttgttct cttttggaat tgttcaaagt 1320

ttttttttta agggatgggt gaatgaactt cagcatccaa gtttatgaat gacagcagtc 1380
acacagttct gtgtatatag ttttaagggt agagtcttgt gttttattca gattgggaaa 1440
tccattctat tttgtgaatt gggataataa cagcagtggg ataagtactt agaaatgtga 1500
aaaatgagca gtaaaataga tgagataaag aactaaagaa attaaagagat agtcaattct 1560
tgctttatac ctgagtttat tctgtaaaat ttttaaagat atatgcatac ctggatttcc 1620
ttggcttctt tgagaatgta agagaaatta aatctgaata aagaattctt cctgttaaaa 1680
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 1722

```

<210> 164
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004988

<400> 164
 cagattggga aatccattct attttgtgaa ttgggataat aacagcagtg gaataagtac 60

<210> 165
 <211> 2334
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004994

<400> 165
 agacacctct gccctcacca tgagcctctg gcagccctcg gtcctgggtgc tccctgggtgct 60
 gggctgctgc tttgctgccc ccagacagcg ccagtcaccc cttgtgtctct tccctggaga 120
 cctgagaacc aatctcaccg acaggcagct ggcagaggaa tacctgtacc gctatgggtta 180
 cactcgggtg gcagagatgc gtggagagtc gaaatctctg gggcctgcgc tgctgcttct 240
 ccagaagcaa ctgtccctgc ccgagaccgg tgagctggat agcgccacgc tgaaggccat 300
 gcgaacccca cgggtgcgggg tcccagacct gggcagattc caaacctttg agggcgacct 360
 caagtggcac caccacaaca tcacctattg gatccaaaac tactcggaag acttgccgcg 420
 ggcgggtgatt gacgacgcct ttgcccgcgc ctctgcactg tggagcgcgg tgacgcgcgt 480
 cacccttact cgcgtgtaca gccgggacgc agacatcgtc atccagtttg gtgtcgcgga 540
 gcacggagac gggatatcct tcgacgggaa ggacgggctc ctggcacacg cctttcctcc 600
 tggccccggc attcaggagg acgcccattt cgacgatgac gagttgtggg ccctgggcaa 660
 gggcgctcgtg gttccaactc ggtttggaaa cgcagatggc ggcgcctgcc acttccccct 720
 catcttcgag ggccgctcct actctgcctg caccaccgac ggtcgctccg acggcttgcc 780
 ctggtgcagt accacggcca actacgacac cgacgaccgg ttggcttctt gccccagcga 840
 gagactctac acccgggacg gcaatgctga tgggaaaccc tgccagtttc cattcatctt 900
 ccaaggccaa tccctactcg cctgcaccac ggacgggtcg tccgacgggt accgctgggt 960
 cgccaccacc gccaaactac accgggacaa gctcttcggc ttctgcccga cccgagctga 1020
 ctcgacgggt atggggggca actcggcggg ggagctgtgc gtcttcccct tcaactttct 1080
 gggtaaggag tactcgacct gtaccagcga gggccgcgga gatgggcgcc tctgggtgcg 1140
 taccacctcg aactttgaca gcgacaagaa gtggggcttc tgcccggacc aaggatacag 1200

```

tttgttcctc gtggcggcgc atgagttcgg ccacgcgctg ggcttagatc attcctcagt 1260
gccggaggcg ctcatgtacc ctatgtaccg cttcaactgag gggcccccct tgcataagga 1320
cgacgtgaat ggcattccggc acctctatgg tcttcgccct gaacctgagc cagggcctcc 1380
aaccaccacc acaccgcagc ccacggctcc cccgacggtc tgccccaccg gacccccac 1440
tgtccacccc tcagagcgcc ccacagctgg ccccaacagg ccccccctcag ctggcccccac 1500
aggtccccc actgctggcc cttctacggc cactactgtg cttttgagtc cgggtggacga 1560
tgcctgcaac gtgaacatct tcgacgccat cgcggagatt gggaaaccagc tgtatttggt 1620
caaggatggg aagtactggc gattctctga gggcaggggg agccggccgc agggccccc 1680
ccttatcgcc gacaagtggc ccgcgctgcc ccgcaagctg gactcgggtct ttgaggagcc 1740
gctctccaag aagcttttct tcttctctgg gcgccagggt tgggtgtaca caggcgcgtc 1800
gggtgctggg ccgaggcgctc tggacaagct gggcctggga gccgacgtgg ccaggtgac 1860
cggggccctc cggagtggca gggggaagat gctgctgttc agcggggcgc gcctctggag 1920

```

```

gttcgacgtg aaggcgcaga tgggtggatcc cccgagcgcc agcgagggtgg accggatggt 1980
ccccgggggt cctttggaca cgcacgacgt cttccagtac cgagagaaag cctatttctg 2040
ccaggaccgc ttctactggc gcgtgagttc ccggagtggc ttgaaccagg tggaccaagt 2100
gggtacgtg acctatgaca tcttcagtg ccttgaggac tagggctccc gtctgtcttt 2160
gcagtgccat gtaaatcccc actgggacca accctgggga aggagccagt ttgccggata 2220
caaactggta ttctgttctg gaggaagggt aggagtggag gtgggctggg cctctctctc 2280
tcacctttgt tttttgttgg agtgtttcta ataaacttgg attctctaac cttt 2334

```

<210> 166

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004994

<400> 166

```

ggccctctct tctcaccttt gttttttggt ggagtgttcc taataaactt ggattctcta 60

```

<210> 167

<211> 5329

<212> DNA

<213> Homo sapiens

<220>

<221> Modified_base

<222> 1 ... 5329

<223> n = a,c,g, or t

<300>

<308> NM_005063

<400> 167

```

gtggtgtcgg tgcgggcagc atccccggcg ccttgcctgg gtgcgcggag ccttcggcct 60
ctgttctcct cccctccccg cccttacctc caccgaggac cgcccgcgcc agtcaactcc 120
tcgcactttg cccctgcttg gcagcggata aaaggggggt gaggaataac cggacacgtc 180
caccgcttgc cagctctagc ctttaaattc ccggtcggg acctccacgc accgggctag 240
cgccgacaac cagctagcgt gcaaggcgcc gcggctcagc gcgtaccggc gggcttcgaa 300
accgcagtcc tccggcgacc ccgaactccg ctccggagcc tcagccccct ggaaagtgat 360
cccggcatcg gagagccaag atgcgggccc acttgcctgca ggacgatata tctagctcct 420
ataccaccac caccaccatt acagcgccct cctccagggt cctgcagaat ggaggagata 480
gatttggagc gatgcccctc tacttgggaag acgacattcg cctgatata aaagatgata 540
tatatgaccc cactacaag gataaggaag gcccaagccc caaggttgaa tatgtctgga 600
gaaacatcat ccttatgtct ctgctacact tgggagccct gtatgggata actttgattc 660
ctacctgcaa gttctacacc tggctttggg gggatttcta ctattttgtc agtgccctgg 720
gcataacagc aggagctcat cgtctgtgga gccaccgctc ttacaaagct cggctgcccc 780
tacggctctt tctgatcatt gccaacacaa tggcattcca gaatgatgtc tatgaatggg 840
ctcgtgacca ccgtgcccac cacaagtttt cagaaacaca tgctgatcct cataattccc 900

```

gacgtggcctt	tttcttctct	cacgtgggtt	ggctgcttgt	gcgcaaacac	ccagctgtca	960
aagagaaggg	gagtacgcta	gacttgctctg	acctagaagc	tgagaaactg	gtgatgttcc	1020
agaggaggta	ctacaaacct	ggcttgctgc	tgatgtgctt	catcctgccc	acgcttgctg	1080
cctgggtattt	ctgggggtgaa	acttttcaaa	acagtgctgtt	cgttgccact	ttcttgcgat	1140
atgctgtggt	gcttaatgcc	acctggctgg	tgaacagtgc	tgcccacctc	ttcggatatc	1200
gtccttatga	caagaacatt	agccccggg	agaatatcct	ggtttcactt	ggagctgtgg	1260
gtgagggcctt	ccacaactac	caccactcct	ttccctatga	ctactctgcc	agtgaagtacc	1320
gctggcacat	caacttcacc	acattcttca	ttgattgcat	ggccgccttc	ggtctggcct	1380
atgaccggaa	gaaagtctcc	aaggccgcca	tcttgccag	gattaaaaga	accggagatg	1440
gaaactacaa	gagtggctga	gtttgggggtc	cctcaggttc	ctttttcaaa	aaccagccag	1500
gcagagggtt	taatgtctgt	ttattaacta	ctgaataatg	ctaccaggat	gctaaagatg	1560
atgatgttaa	cccattccag	tacagtattc	ttttaaaatt	caaaagtatt	gaaagccaac	1620
aactctgcct	ttatgatgct	aagctgatat	tatttcttct	cttatcctct	ctctcttcta	1680
ggcccatgt	cctccttttc	actttaatcg	ccctcctttc	ccttattgcc	tcccaggcaa	1740
gcagctggtc	agtctttgct	cagtgtccag	cttccaaagc	ctagacaacc	tttctgtagc	1800
ctaaaacgaa	tggtctttgc	tccagataac	tctctttcct	tgagctgttg	tgagctttga	1860
agtaggtggc	ttgagctaga	gataaaacag	aatcttctgg	gtagtccct	gttgattatc	1920
ttcagcccag	gcttttgcta	gatggaatgg	aaaagcaact	tcatttgaca	caaagcttct	1980
aaagcnaggt	aaattgtcgg	gggagagagt	tagcatgtat	gaatgtaagg	atgagggaag	2040
cgaaggaacc	tctcgccatg	atcagacata	cagctgccta	cctaattgag	acttcaagcc	2100
ccaccacata	gcatgcttcc	tttctctcct	ggctcggggg	aaaaagtggc	tgccggtgtt	2160
ggcaatgcta	attcaatgcc	gcaacatata	gttgaggccg	aggataaaga	aaagacattt	2220
taagtittga	gtaaaaagtgg	tctctgctgg	ggaagggttt	tcttttcttt	ttttctttta	2280
taacaaggag	atttcttagt	tcatatatca	agaagctctg	aagttgggtg	tttccagaat	2340
tggtaaaaac	agcagctcat	agaattttga	gtattccatg	agctgctcat	tacagtctct	2400
tcctctttct	gctctgccat	cttcaggata	ttgggtcttc	ccctcatagt	aataagatgg	2460
ctgtggcatt	tccaaacatc	caaaaaaagg	gaaggattta	aggaggtgaa	gtccgggtcaa	2520
aaataaaaata	tatatacata	tatacattgc	ttagaacgtt	aaactattag	agtatttccc	2580
ttccaaagag	ggatgtttgg	aaaaaactct	gaaggagagg	aggaattagt	tgggatgcc	2640
atttctctct	cactgctgga	catgagatgg	agaggctgag	ggaccaggatc	tataggcagc	2700
ttctaagagc	gaacttcaca	taggaaggga	tctgagaaca	cgttcagggg	ttgagaagg	2760
tactgagtga	gttattggga	gtcttaataa	actagatatt	aggtccattc	attaattagt	2820
tccagtttct	ccttgaaatg	agtaaaaact	agaaggcttc	tctccacagt	gttgtgcccc	2880
ttcactcatt	tttttttgag	gagaaggggg	tctctgttaa	catctagcct	aaagtataca	2940
aactgcctgg	ggggcagggt	taggaatctc	ttcactacct	tgattcttga	ttcctggctc	3000
taccctgtct	gtcccccttc	tttgaccaga	tctttctctt	ccctgaacgt	tttcttcttt	3060
ccctggacag	gcagcctcct	ttgtgtgtat	tcagaggcag	tgatgacttg	ctgtccaggc	3120
agctccctcc	tgcacacaga	atgctcagg	tcactgaacc	actgcttctc	ttttgaaagt	3180
agagctagct	gccactttca	cgtggcctcc	gcagtgtctc	cacctacacc	cctgtgctcc	3240
cctgccacac	tgatggctca	agacaaggct	ggcaaacctc	cccagaaaca	tctctggccc	3300
agaaagcctc	tctctccctc	cctctctcat	gagaagccaa	gcgctcatgt	tgagccagtg	3360
ggccagccac	agagcaaaaag	agggtttatt	ttcagtcctc	tctctctggg	tcagaaccag	3420
agggcatgct	gaatgcccc	tgcttacttg	gtgagggtgc	ccgcctgag	tcagtgtctc	3480
cagctggcag	tgcaatgctt	gtagaagtga	gaggaaacag	ttctcactgg	gaagaagcaa	3540
gggcaagaac	ccaagtgcct	cacctcgaaa	ggaggccctg	ttccctggag	tcagggtgaa	3600
ctgcaaagct	ttggctgaga	cctgggattt	gagataccac	aaacctgct	gaacacagtg	3660
tctgttcagc	aaactaacca	gcattcccta	cagcctaggg	cagacaatag	tatagaagtc	3720
tggaaaaaaa	caaaaacaga	atttgagaac	cttggaccac	tcctgtccct	gtagctcagt	3780
catcaaagca	gaagtctggc	tttgctctat	taagattgga	aatgtacact	accaaact	3840
cagtccactg	ttgagcccca	gtgctggaag	ggagggaaggc	ctttcttctg	tgtttaattgc	3900
gtagaggcta	caggggttag	cctggactaa	aggcatcctt	gtctttgagc	tattcacctc	3960
agtagaaaag	gatctaagg	aagatcactg	tagtttagtt	ctgttgacct	gtgcacctac	4020
cccttgaaa	tgtctgctgg	tatttctaat	tccacaggtc	atcagatgcc	tgcttgataa	4080
tatataaaca	ataaaaacaa	ctttcacttc	ttcctattgt	aatogtgtgc	catggatctg	4140
atctgtacca	tgacctaca	taaggctgga	tggcacctca	ggctgagggc	cccaatgtat	4200
gtgtggctgt	gggtgtgggt	gtgctgagta	aggaacacga	aggaacacga	ttttcaagat	4260
tctaaagctc	aattcaagt	acacattaat	gataaactca	gatctgatca	agagtccgga	4320
tttctaacag	tccttgcttt	gggggtgtg	ctggcaactt	agctcagggtg	ccttacatct	4380
tttctaataca	cagtgttgca	tatgagcctg	ccctcactcc	ctctgcagaa	tccctttgca	4440
cctgagaccc	tactgaagt	gctggtagaa	aaaggggcct	gagtggagga	ttatcagtat	4500
cacgatttgc	aggattccct	tctgggcttc	attctggaaa	cttttggttag	ggctgctttt	4560

```

cttaagtgcc cacatttgat ggaggggtgga aataatttga atgtatttga tttataagtt 4620
tttttttttt tttgggttaa aagatgggttg tagcatttaa aatggaaaat tttctccttg 4680
gtttgctagt atcttgggtg tattctctgt aagtgtagct caaatagggtc atcatgaaag 4740
gttaaaaaag cgaggtggcc atgttatgct ggtgggttgc agggcctcca accactgtgc 4800
cactgacttg ctgtgtgacc ctgggcaagt cacttaacta taagggtgcct cagttttcct 4860
tctgttaaaa tgggggataat aatactgacc tacctcaaag ggcagttttg aggcagtact 4920
aatgcttttt agaaagcatt ttgggatcct tcagcacagg aattctcaag acctgagtat 4980
tttttataat aggaatgtcc accatgaact tgatacgtcc gtgtgtccca gatgctgtca 5040
ttagtctata tggttctcca agaaactgaa tgaatccatt ggagaagcgg tggataacta 5100
gccagacaaa atttgagaat acataaacia cgcattgcca cggaacata cagaggatgc 5160
cttttctgtg attgggtggg attttttccc tttttatgtg ggatatagta gttacttgtg 5220
acaagaataa ttttggaata atttctatta atatcaactc tgaagctaatt tgtactaatc 5280
tgagattgtg tttgttcata ataaaagtga agtgaatctg attgcactg 5329

```

<210> 168
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005063

<400> 168
 aataatgcta ccaggatgct aaagatgatg atgttaaccc attccagtac agtattcttt 60

<210> 169
 <211> 634
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005101

```

<400> 169
cggctgagag gcagcgaact catcttttgc agtacaggag cttgtgccgt ggcccacagc 60
ccacagccca cagccatggg ctgggacctg acggtgaaga tgctggcggg caacgaattc 120
caggtgtccc tgagcagctc catgtcgggtg tcagagctga aggcgcagat caccagaag 180
attggcgtgc acgccttcca gcagcgtctg gctgtccacc cgagcgggtg ggcgctgcag 240
gacaggggtc cccttgccag ccagggcctg ggccctggca gcacggctct gctgggtggg 300
gacaaatgcg acgaacctct gagcatcctg gtgaggaata acaagggccg cagcagcacc 360
tacgaggtcc ggctgacgca gaccgtggcc cacctgaagc agcaagtgcg cgggctggag 420
ggtgtgcagg acgacctgtt ctggctgacc ttcgagggga agcccttggg ggaccagctc 480
ccgctggggg agtacggcct caagcccctg agcacctgtg tcatgaatct gcgcctgcgg 540
ggaggcggca cagagcctgg cgggcgggag taagggcctc caccagcatc cgagcaggat 600
caagggccgg aaataaaggc tgttgtaaga gaat 634

```

<210> 170
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005101

<400> 170
 tgggtgtgga caaatgcgac gaacctctga gcatcctggt gaggaataac aagggccgca 60

<210> 171
 <211> 1339
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_005139

<400> 171

```

gaattccgat tagtgtgata tcagctcaag gcaaagggtgg gatatcatgg catctatctg 60
ggttggacac cgaggaacag taagagatta tccagacttt agcccatcag tggatgctga 120
agctattcag aaagcaatca gaggaattgg aactgatgag aaaatgctca tcagcattct 180
gactgagagg tcaaatgcac agcggcagct gattgttaag gaatatcaag cagcatatgg 240
aaaggagctg aaagatgact tgaagggtga tctctctggc cactttgagc atctcatggt 300
ggccctagtg actccaccag cagtctttga tgcaaagcag cttaaagaaat ccatgaaggg 360
cgcgggaaca aacgaagatg ccttgattga aatcttaact accaggacaa gcaggcaaat 420
gaaggatata tctcaagcct attatacagt atacaagaag agtcttgagg atgacattag 480
ttccgaaaca tctggtgact tccggaaago tctgttgact ttggcagatg gcagaagaga 540
tgaaagtctg aaagtggatg agcatctggc caaacaagat gccagattc tctataaagc 600
tgggtgagaac agatggggca cggatgaaga caaattcact gagatcctgt gtttaaggag 660
ctttcctcaa ttaaaactaa catttgatga atacagaaat atcagccaaa aggacattgt 720
ggacagcata aaaggagaat tatctgggca ttttgaagac ttactgttgg ccatagttaa 780
ttgtgtgagg aacacgccgg ccttttttagc cgaaagactg catcgagcct tgaagggtat 840
tggaactgat gagtttactc tgaaccgaat aatggtgtcc agatcagaaa ttgacctttt 900
ggacattcga acagagttca agaagcatta tggctattcc ctatattcag caattaaatc 960
ggatacttct ggagactatg aaatcacact cttaaaaaatc tgtggtggag atgactgaac 1020
caagaagata atctccaaag gtccacgatg ggctttccca acagctccac cttacttctt 1080
ctcatactat ttaagagaac aagcaaatat aaacagcaac ttgtgttctt aacaggaatt 1140
ttcattgttc tataacaaca acaacaaaag cgattattat tttagagcat ctcatttata 1200
atgtagcagc tcataaatga aattgaaaat ggtattaaag atctgcaact actatccaac 1260
ttatatttct gctttcaaag ttaagaatct ttatagttct actccattaa atataaagca 1320
agataataaa acggaattc 1339

```

<210> 172

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005139

<400> 172

```

ttcagcaatt aaatcggata cttctggaga ctatgaaatc acactcttaa aaatctgtgg 60

```

<210> 173

<211> 1582

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005165

<400> 173

```

ccgagctgtg cttgtggctg cggctgctaa ctggctgcgc acagggagct gtcaccatgc 60
ctcactcgta cccagccctt tctgctgagc agaagaagga gttgtctgac attgccctgc 120
ggattgtagc cccgggcaaa ggcattcttg ctgaggatga gtctgtaggc agcatggcca 180
agcggctgag ccaaattggg gtggaaaaca cagaggagaa ccgccggctg taccgccagg 240
tcctgttcag tgctgatgac cgtgtgaaaa agtgcattgg aggcgtcatt ttcttccatg 300
agacctcta ccagaaagat gataatggtg ttcccttcgt ccgaaccatc caggataagg 360
gcatcgctgt gggcatcaag gttgacaagg gtgtggtgcc tctagctggg actgatggag 420
aaaccaccac tcaagggtct gatgggctct cagaacgctg tgccaatac aagaaggatg 480
gtgctgactt tgccaagtgg cgctgtgtgc tgaaaatcag tgagcgtaca ccctctgcac 540
ttgccattct ggagaacgcc aacgtgctgg ccggttatgc cagtatctgc cagcagaatg 600
gcattgtgac tattgtggaa cctgaaatat tgctgatgg agaccacgac ctcaaacggt 660
gtcagtatgt tacagagaag gtcttggctg ctgtgtacaa ggccctgagt gaccatcatg 720

```

```

tatacctgga ggggaccctg ctcaagccca acatgggtgac cccgggcat gcctgtccca 780
tcaagtatac cccagaggag attgccatgg caactgtcac tgccctgcgt cgcactgtgc 840
ccccagctgt cccaggagtg accttcctgt ctgggggtca gagcgaagaa gaggcacatcat 900
tcaacctcaa tgccatcaac cgctgcccc ttccccgacc ctgggcgctt accttctcct 960
atgggcgtgc cctgcaagcc tctgcaactca atgcctggcg agggcaacgg gacaatgctg 1020
gggctgccac tgaggagtgc atcaagcggg ctgaggtgaa tgggcttgca gccagggca 1080
agtatgaagg cagtggagaa gatgggtggag cagcagcaca gtcactctac attgccaacc 1140
atgcctactg agtatccact ccataccaca gcccttggcc cagccatctg caccacttt 1200
tgctttagt catggccagg gccaaatagc tatgcagagc agagatgcct tcacctggca 1260
ccaacttgtc ttcccttctc tcttcccttc cctctctca ttgctgcacc tgggaccata 1320
ggatgggagg atagggagcc cctcatgact gagggcagaa gaaattgcta gaagtcagaa 1380
caggatggct gggctctccc ctacctcttc cagctccac aattttccca tgatgaggta 1440
gcttctccct gggtctctct tcttgccctgc cctgtctcct gggatcagag ggtagtagag 1500
aagccctgac tcatgccttg agtacatacc atacagcaaa taaatggtag caaaacaaaa 1560
aaaaaaaaa aaaaaaaaaa aa 1582

```

<210> 174

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005165

<400> 174

gagggtagta cagaagccct gactcatgcc ttgagtacat accatacagc aaataaatgg 60

<210> 175

<211> 451

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005213

<400> 175

```

acttccctgt tcaactttggt tccagcatcc tgtccagcaa agaagcaatc agccaaaatg 60
atacctggag gcttatctga ggccaaaccc gccactccag aaatccagga gattgttgat 120
aagggttaaac cacagcttga agaaaaaaca aatgagactt atggaaaatt ggaagctgtg 180
cagtataaaa ctcaagttgt tgctggaaca aattactaca ttaaggtagc agcaggtgat 240
aataaatata tgcacttgaa agtattcaaa agtcttcccg gacaaaatga ggacttggtg 300
cttactggat accagggttga caaaaacaag gatgacgagc tgacgggctt ttagcagcat 360
gtacccaaag tggtctgatt ccttcaactg gctactgagt catgatcctt gctgataaat 420
ataaccatca ataaagaagc attcttttcc a 451

```

<210> 176

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005213

<400> 176

aactggctac tgagtcatga tccttgctga taaatataac catcaataaa gaagcattot 60

<210> 177

<211> 366

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005218

<400> 177

```

gtcagctcag cctccaaagg agccagcctc tccccagttc ctgaaatcct gagtgttgcc 60
tgccagtcgc catgagaact tectaccttc tgctgtttac tctctgctta cttttgtctg 120
agatggcctc aggtggtaac tttctcacag gccttggcca cagatctgat cattacaatt 180
gcgtcagcag tggagggcaa tgtctctatt ctgcctgccc gatctttacc aaaattcaag 240
gcacctgtta cagaggggaag gccaaagtgt gcaagtgagc tgggagtgac cagaagaaat 300
gacgcagaag tgaaatgaac tttttataag cattctttta ataaaggaaa attgcttttg 360
aagtat 366

```

<210> 178

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005218

<400> 178

```

gggagtgacc agaagaaatg acgcagaagt gaaatgaact ttttataagc attcttttaa 60

```

<210> 179

<211> 1519

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005326

<400> 179

```

ctgcctcggg acgctgtccc ccgcagcgac ggcccgttcc acctcgcgat ctgccgggta 60
cccgggcggg gtggcgctcg gcctccaggg atccactgtg cggtgccaaa aaagaggcgg 120
aggctcgcgg cacagctctc ccggcgcgac tctcggggcg ccgccgcgcg tcccaggccc 180
gtctcccggc ccgtggcagt cggggctcgc ggacaaaaca agttgagcgc gagcgcgctg 240
attggtttggc ggacgggtgcg aggtggacgc tgattggctg agggcagcgc gagggcgggcg 300
ctgattggct gcgacgcgcc gacgccggtg ttttgcagtc ctgggcagct cggcagtgca 360
gcccggcccc ggtcatgggt gtgggcccag ggctgtcctg ccgccgcagc ctccgcgcgc 420
tgggagccgc ctgcgcccgc cgaggcctcg gtccagccct gctgggagtt ttctgccaca 480
cagatttgcg gaagaacctg accgtggacg agggcaccat gaaggtagag gtgctgcctg 540
ccctgaccga caactacatg tacctggtca ttgatgatga gaccaaggag gctgccattg 600
tgatcccggt gcagccccag aaggctcgtg acgcggcgag aaagcacggg gtgaaactga 660
ccacagtgtc caccaccac caccactggg accatgctgg cgggaatgag aaactgggtca 720
agctggagtc gggactgaag gtgtacgggg gtgacgaccg tatcggggcc ctgactcaca 780
agatcactca cctgtccaca ctgcaggtgg ggtctctgaa cgtcaagtgc ctggcgaccc 840
cgtgccacac ttcaggacac atttgttact tctgtagcaa gcccggaggc tcggagcccc 900
ctgccgtgtt cacaggtgac acctgtttg tggtggctg cgggaagtgc tatgaaggga 960
ctgcggatga gatgtgtaaa gctctgctgg aggtcttggg ccggtctccc ccggacacaa 1020
gagtctactg tggccacgag tacaccatca acaacctcaa gtttgacgc cacgtggagc 1080
ccggcaatgc cgccatccgg gagaagctgg cctgggcca gagaagtac agcatcgggg 1140
agccacagtc gccatccacc ctggcagagg agtttaccta caacccttc atgagagtga 1200
gggagaagac ggtgcagcag cacgcaggtg agacggaccc ggtgaccacc atgctgggccg 1260
tgcgcaggga gaaggaccag ttcaagatgc cccgggactg aggccgccct gcaccttcag 1320
cggatttggg gattaggtc ttttagttaa ctggctttcc tgctgggtccg tgcgggaaat 1380
tcagtcttga tttaacctta attttacagc ccttggcttg tgttatcgga cattctaagt 1440
catatttata agagaagttt aacaagtatt tattcccata aaaaaaaaaa aaaaaaaaaa 1500
aaaaaaaaaa aaaaaaaaaa 1519

```

<210> 180

<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_005326

<400> 180
cttgtgttat cggacattct aatgcatatt tataagagaa gtttaacaag tattttattcc 60

<210> 181
<211> 3378
<212> DNA
<213> Homo sapiens

<300>
<308> NM_005461

<400> 181
acagctgcac cgccgagctg cgagcggctg cgagcgagag agcgttaagag caagagagct 60
agagagcgag caacggggcac tcgccccacg cctccccctca gccccaccgc gcgctccgct 120
tgctctccca ccccgcccca ctctaccccg cccggctccct gcgcgggcac agcccagagc 180
tctggggcgg tgcaggcagc ctccgggactc tcgggcgcgc cgcgcgctcc ccagacaaag 240
gcttgggcgg cgcccccgcc ccgctgcgcc ctccgtcccc gcctccccag ctcttctccg 300
ctcttcccc ccgcgcttgg ctccggcggc tcgggcccgc cgcaaagtct cccgggcggc 360
agcggcggct gcgcctcgct tcagcgatgg ccgcgagct gagcatgggg ccagagctgc 420
ccaccagccc gctggccatg gactatgtca acgacttcga cctgctcaag ttcgacgtga 480
agaaggagcc actggggcgc gcggagcgct cgggcaggcc ctgcacacgc ctgcagccag 540
ccggctcggg gtcctccaca ccgctcagca ctccgtgtag ctccgtgccc tcgtcgccca 600
gcttcagccc gaccgaacag aagacacacc tgagggatct tcgaggatct gtactggatg gcgagcaact 660
accagcagat gaaccccgag gcgctcaacc tgacgcccga ggacgcggtg gaagcgctca 720
tcggctcgca cccagtgcga cagccgctgc aaagcttcga cagctttcgc ggcgctcacc 780
accaccacca tcaccaccac cctcaccgcg accacgcgta cccgggcgccc ggcgtggccc 840
acgacgagct gggcccgcac gctcaccgcg accatcacca tcatcaccaa gcgtcgccc 900
cgccgtccag cgccgctagc ccggcgcaac agctgcccac tagccacccc gggcccgggc 960
cgcacgcgac ggcctcggcg acggcgggcg gcggcaacgg cagcgtggag gaccgcttct 1020
ccgaogacca gctcgtgtcc atgtccgtgc gcgagctgaa ccgccacctg cggggcttca 1080
ccaaggacga ggtgatccgc ctgaagcaga agcggcgag cctgaagaac cggggctacg 1140
cccagctctg caggtataaa cgcgtccagc agaagcacca cctggagaat gagaagacgc 1200
agctcattca gcaggtggag cagcttaagc aggaggtgtc ccggctggcc cgcgagagag 1260
acgcctacaa ggtcaagtgc gaaaaactcg ccaactccgg cttcagggag gcgggctcca 1320
ccagcgacag cccctcctct cccgagttct ttctgtgagt cgtggccggg cctggcccc 1380
gcccttgccc cgccccggac tccctgtccc acgtccctag tcccagacta ccccgacccc 1440
tgtccctgccc gcggccccag ccttgacctg tttgacttga gcgagagggg ggaagggcgc 1500
gcgggcccgc ggcgacgggc ggggtgcgcg gcgggcaggg gaccttggct aaggcgagag 1560
tagcgcacgc cagcgccgcc tccctagact gagcagagcc ggagagagag acgagagggg 1620
gggaggtccc ggagtaactt ctctccaggg tgaagggcgg cgaggcatag tcccgagaag 1680
tcaccaaggc catctggaga ctcttggtt tctgaacttt gcgcgttaag ccgggacagc 1740
tgctttgctg cccggagagt agtccgcgcg aggaagagag caacgaggaa aggagagggg 1800
ctctggcgct ccggcaggcg agaggcgagg ctgagcgaaa gaaggaagga cagacggacc 1860
tgtctgtcag agttcggaga acactggctc tcagccctga gacacaggcc tcagtttagg 1920
cgctcggcgc ccaaactctc tcagttttat tgctgtctcg attatataga aaaatacaaa 1980
aaatctgcat taaaaatatt aatcctgcat gctggacatg tatggtaata atttctattt 2040
tgtaccattt tcttgtttta ctttagcatg ttgttgatca tggatcatac tcccctggt 2100
tctttgggtg agaagggatc gcagtttgga aactccggcg gctcgtgcg ggggttctagt 2160
cccagctgta ggcttgtaaa taccgcccc gccaaacgc atagagaacg tggcagcaag 2220
ctgagggctc ttgtttgggt ttattattac ggtatttttg tttgtaagtt aaaaagaaaa 2280
aaaaaaagaa aaagtctcgg gcattttgca tcagaaaaca actttgtctt ggggcacact 2340
tggaagtgtc atgttttctt tccttccctt atccccattc ggtcctcttt ttcctctctc 2400
gcttttagttt tcaaccttgt tgggtgctgag agagagaacc gagaggtccc agtacaaggg 2460
cagggcaggg caggggaagc gccaaagctc gcaccccaga ggagtgttct ggactacagc 2520
cttgtcttat ggtcaaattg atacccttaa taagaaagga aaggaaagga aaacagatcc 2580

```

tccccctctgc tttttattgt aaccagaatc accctgaggt cccttctgaa ccctctgggc 2640
ctgcgctaata tgtaggagcc acagcgctcc taggggtgaga ggcttagcca tccctgaccc 2700
tggcagtgca ctggtaagca gacactgcac tgaaccaact gctatgctca gaatgtacca 2760
gaaacccaaa cattggcaag taattttgca actttcaagt gcgttcttta gaccaatgca 2820
ttgogtttct ttccctgctt ttgagatagt aggaagagtt cttgggtggtg tccccccct 2880
tcaattcttc agttgtatag tagttatagg gaagatatgg gtgtttttct ttattattac 2940
tttttttttt ctgcaggtca gtaaaaggat ttaagttgca ctgacaaaaa taccaaaata 3000
aaagtgtatt ttttaagttcc catttgaaat tgctggcgct gctggccgga tgcatttttg 3060
agtttgtatt agttgataaa ttaacagtaa taacaagatt gtatgaaccg catgggtgctt 3120
gcagttttta atattgtgga tatttgtcct gcatcagaaa cgagctttgg tttttacaga 3180
ttcaactgtg ttgaaatcaa acctgccgca acagaaattg tttttatttc atgtaaaata 3240
agggatcaat ttcaaaccct gcttatgata tgaaaatatt aaaacctagt ctattgtagt 3300
tttattcaga ctggtttctg ttttttggtt attaaaatgg tttcctattt tgcttattaa 3360
aaaaaaaaaa aaaaaaaaa 3378

```

<210> 182
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005461

<400> 182
 atttgtcctg catcagaaac gagctttggt ttttacagat tcaactgtgt tgaaatcaaa 60

<210> 183
 <211> 597
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005532

<400> 183
 agctgaagtt gaggatctct tactctctaa gccacggaat taaccgcgagc aggcattggag 60
 gcctctgctc tcacctcatc agcagtgacc agtggtggcca aagtgggtcag ggtggcctct 120
 ggctctgccc tagttttgcc cctggccagg attgctacag ttgtgattgg aggagttgtg 180
 gccatggcgg ctgtgcccat ggtgctcagt gccatgggct tcaactgccc gggaatcgcc 240
 tcgtcctcca tagcagccaa gatgatgtcc ggggcggcca ttgccaatgg ggggtggagt 300
 gcctcgggca gccttggtgg tactctgcag tcaactgggag caactggact ctccggattg 360
 accaagttca tcctgggctc cattgggtct gccattgcgg ctgtcattgc gaggttctac 420
 tagctcctg ccctcgccc tgcagagaag agaaccatgc caggggagaa ggcacccagc 480
 catcctgacc cagcaggag ccaactatcc caaatatacc tgggtgaaat ataccaaatt 540
 ctgcatctcc agaggaaaat aagaaataaa gatgaattgt tgcaactctt aaaaaaa 597

<210> 184
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005532

<400> 184
 agccaactat cccaaatata cctgggtgaa atataccaaa ttctgcatct ccagaggaaa 60

<210> 185
 <211> 1661
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_005566

<400> 185

```

tgctgcagcc gctgccgccc attccggatc tcattgccac ggcgccccga cgaccgcccc 60
acgtgcattc ccgattccctt ttgggttccaa gtccaatatg gcaactctaa aggatcagct 120
gatttataat cttctaaagg aagaacagac cccccagaat aagattacag ttgttgggg 180
tggtgctggt ggcatggcct gtgccatcag tatcttaatg aaggacttgg cagatgaact 240
tgctccttgtt gatgtcatcg aagacaaatt gaaggagag atgatggatc tccaacatgg 300
cagccttttc cttagaacac caaagattgt ctctggcaaa gactataatg taactgcaaa 360
ctccaagctg gtcattatca cggctggggc acgtcagcaa gagggagaaa gccgtcttaa 420
tttgggtccag cgtaacgtga acatatttaa attcatcatt cctaattgtt taaaatacag 480
cccgaactgc aagttgctta ttgtttcaaa tccagtggat atcttgacct acgtggcttg 540
gaagataagt ggttttccca aaaaccgtgt tattggaagt ggttgcaatc tggattcagc 600
ccgattccgt tacctgatgg gggaaaggct gggagttcac ccattaagct gtcattgggtg 660
ggtccttggg gaacatggag attccagtgt gcctgtatgg agtggaatga atgttgctgg 720
tgtctctctg aagactctgc acccagattt agggactgat aaagataagg aacagtggaa 780
agaggttcac aagcaggtgg ttgagagtgc ttatgaggtg atcaaactca aaggctacac 840
atcctgggct attggactct ctgtagcaga tttggcagag agtataatga agaactctag 900
gcgggtgcac ccagtttcca ccatgattaa gggctctttac ggaataaagg atgatgtctt 960
ccttagtggt ccttgcatth ttggacagaa tggaaatctc gaccttgtga aggtgactct 1020
gacttctgag gaagaggccc gtttgaagaa gagtgcagat acactttggg ggatccaaaa 1080
ggagctgcaa ttttaaagtc ttctgatgtc atatcatttc actgtctagg ctacaacagg 1140
attctaggtg gaggttggtc atgttgctct ttttatctga tctgtgatta aagcagtaat 1200
attttaagat ggactgggaa aaacatcaac toctgaagtt agaaataaga atgggttgta 1260
aaatccacag ctatatcctg atgctggatg gtattaatct tgtgtagtct tcaactgggt 1320
agtgtgaaat agttctgcca cctctgacgc accactgcca atgctgtacg tactgcattt 1380
gccccttgag ccaggtggat gtttaccgtg tgttatataa ctctctggct cttcactga 1440
acatgcctag tccaacattt tttcccagtg agtcacatcc tgggatccag tgtataaatc 1500
caatatcatg tcttgtgcat aattcttcca aaggatctta ttttgtgaac tatatcagta 1560
gtgtacatta ccatataatg taaaaagatc tacatacaaa caatgcaacc aactatccaa 1620
gtgttatacc aactaaaacc cccaataaac cttgaacagt g 1661

```

<210> 186

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005566

<400> 186

```

catcaactcc tgaagttaga aataagaatg gtttgtaaaa tccacagcta tatcctgatg 60

```

<210> 187

<211> 2993

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005689

<400> 187

```

gggcctgcag ttggcagaag ggtccccggc ccagagccag cggggccgtg ctgagacggc 60
gtacgtgccc tgcgtgagtg cgtggcggcg gcgcgtgcgc taggggagtg ggcgggtgagg 120
cctgggtccac gtgcgtccct tcccgggacc cccgcagctt ggcgccagc ggctacgtga 180
gccaaggcac ccgatgtcc gcgcccctct ccgagtgaac agtcccggcc tccgggtccc 240
cagtccccgc agcctcggcc ggcgtccacg cattgccatg gtgactgtgg gcaactactg 300
cgaggccgaa gggcccgtgg gtccggcctg gatgcaggat ggccctgagtc cctgcttctt 360
cttcacgctc gtgccctcga cgcggatggc tctagggact ctggccttgg tgctggctct 420

```

```

ccctgcaga cgccgggagc ggcccgtctg tgctgattcg ctgtcttggg gggccggccc 480
gcacatctct cctacgtgc tgcagctgct tctggccaca cttcaggcgg cgctgcccct 540
ccgggcttg gctggccggg tgggactgct ccggggggcc cactgccaa gctatctact 600
tgggctcc gtgctggaga gtctggccgg cgctgtggc ctgtggctgc ttgtcgtgga 660
ggagccag gcacggcagc gtctggcaat gggcatctgg atcaagttca ggcacagccc 720
gtctctctg ctctcttgga ctgtggcggt tgcagctgag aacttggccc tgggtgtcttg 780
acagccca cagtgggtgg gggcaagggc agacttgggc caacagggtt agtttagcct 840
gggtgctg cggatgtgg tctctggagg gctgtttgtc ctgggtctct gggcccctgg 900
ttcgtccc cagtccctata cattgcagggt tcatgaagag gaccaagatg tggaaaggag 960
aggttcgg tcagcagccc aacagtctac ctggcgagat tttggcagga agctccgcct 1020
tgagtggc tacctgtggc ctcgaggag tccagctctg cagctggtgg tgctcatctg 1080
tggggctc atgggtttgg aacgggcaact caatgtgttg gtgcctatat tctataggaa 1140
ttgtgaac ttgtgactg agaaggcacc ttggaactct ctggcctgga ctgttaccag 1200
acgtcttc ctcaagttcc tccagggggg tggcactggc agtacaggct tcgtgagcaa 1260
tgcgccac ttctcttgga tccgggtgca gcagttcacg tctcggcggg tggagctgct 1320
tcttctcc cactctcacg agctctcact gcgctggcac ctggggcgcc gcacagggga 1380
tgctgcgg atcgcgatc ggggcacatc cagtgtcaca gggctgctca gctacctggt 1440
tcaatgtc atccccacgc tggccgacat catcattggc atcatctact tcagcatgtt 1500
tcaacgcc tggtttggcc tcattgtgtt cctgtgcatg agtctttacc tcaccctgac 1560
ttgtgggtc actgagtgga gaaccaagtt tcgtcgtgct atgaacacac aggagaacgc 1620
cccgggca cgagcagtggt actctctgct aaacttcag acggtgaagt attacaacgc 1680
tagagttac gaagtggaa gctatcgaga ggccatcacc aaatatcagg gtttggagtg 1740
tagtcgagc gcttcaactg ttttactaaa tcagaccag aacctggtga ttgggctcgg 1800
tctctgcc ggctccctgc tttgcgcata ctttgtcact gagcagaagc tacagggttg 1860
tactatgtg ctcttttgga cctacattat ccagctgtac atgcccctca attgggtttg 1920
tctactac aggatgatcc agaccaactt cattgacatg gagaacatgt ttgacttgct 1980
aagaggag acagaagtga aggaccttc tggagcaggg ccccttcgct ttcagaaggg 2040
gtatttgag tttgagaacg tgcacttcag ctatgccgat gggcgggaga ctctgcagga 2100
tgctcttc actgtgatgc ctggacagac acttgccctg gtgggcccac ctggggcagg 2160
tagagcac attttgcgc tgetgtttcg cttctacgac atcagctctg gctgcatccg 2220
tagatggg caggacattt cacagggtgac ccaggcctct ctcgggtctc acattggagt 2280
tgccccaa gacactgtcc tctttaatga caccatcgcc gacaatatcc gttacggccg 2340
tcacagct gggaatgat aggtggaggc tgctgctcag gctgcaggca tccatgatgc 2400
ttatgggt tccctgaag ggtacaggac acagggtgggc gagcggggac tgaagctgag 2460
tgccgggag aagcagcgcg tcgccattgc ccgcaccatc ctcaaggctc cgggcatcat 2520
tgctggat gaggcaacgt cagcgctgga tacatctaata gagagggcca tccaggcttc 2580
tgggccaaa gtctgtgcca accgcaccac catcgtagtg gcacacaggc tctcaactgt 2640
tcaatgct gaccagatcc tcgtcatcaa ggatggctgc atcgtggaga ggggacgaca 2700
taggctctg ttgtcccag gtgggtgtga tgctgacatg tggcagctgc agcagggaca 2760
gaagaaacc tctgaagaca ctaagcctca gaccatggaa cgggtgacaaa agtttggcca 2820
tccctctc aaagactaac ccagaaggga ataagatgtg tctcctttcc ctggcttatt 2880
atcctgggt cttgggttat ggtgctagct atggtaaggg aaagggacct ttccgaaaaa 2940
atcttttgg ggaaataaaa atgtggactg tgaaaaaaa aaaaaaaa aaa 2993

```

```

10> 188
11> 60
12> DNA
13> Homo sapiens

```

```

100>
108> NM_005689

```

```

100> 188
jaaagggac ctttccgaaa aacatctttt ggggaaataa aaatgtggac tgtgaaaaaa 60

```

```

10> 189
11> 1830
12> DNA
13> Homo sapiens

```

```

100>

```

<308> NM_005749

<400> 189

```

ggggagttga aacctaatTT tgtggcgtag cagctatgca gcttgaaatc caagtagcac 60
taaattttat tatttcgtat ttgtacaata agcttcccag gagacgtgtc aacatttttg 120
gtgaagaact tgaaagactt cttaagaaga aatatgaagg gcactgggtat cctgaaaagc 180
catacaaagg atcgggggtt agatgtatac acatagggga gaaagtggac ccagtgattg 240
aacaagcatc caaagagagt ggtttggaca ttgatgatgt tcgtggcaat ctgccacagg 300
atcttagtgt ttggatcgac ccatttgagg ttctttacca aattgggtgaa aagggaccag 360
tgaagggtgt ttacgtggat gataataatg aaaatggatg tgagttggat aaggagatca 420
aaaacagctt taaccacagag gcccagggtt ttatgcccac aagtgaacca gcctcatcag 480
tgtccagctc tccatgcgct ccttttgggt actctgctgc tgtaagccct accttcatgc 540
cccgggtccac tcagccttta acctttacca ctgccacttt tgctgccacc aagttcgggt 600
ctacaaaaat gaagaatagt ggccgtagca acaagggtgc acgtacttct cccatcaacc 660
tcggcttgaa tgtgaatgac ctcttgaaagc agaaagccat ctcttctctca atgcactctc 720
tgtatgggct tggcttgggt agccagcagc agccacagca acagcagcag ccagcccagc 780

cgccaccgcc accaccacca ccacagcagc aacaacagca gaaaacctct gctctttctc 840
ctaattgcaa ggaatttatt ttctctaata tgcagggtca aggtagtagt accaatggaa 900
tgttcccagg tgacagcccc cttaacctca gtctctcca gtacagtaat gcctttgatg 960
tgtttgcagc ctatggaggg ctcaatgaga agtcttttgt agatgggttg aattttagct 1020
taaataacat gcagtattct aaccagcaat tccagcctgt tatggctaac taaaaaaaag 1080
aaaatgtatc gtacaagtta aaatgcacgg gcccaagggg gatttttttt ttcacctcct 1140
tgagaatttt tttttttaag cttatagtaa ggatacatte aagottgggt aaaaaaataa 1200
taataaaaca tgcattcttt ttcatattgc aaccaagcac aaagtatttt tatactgact 1260
gtatatttta aagtatactc tcagatatgg cctcttacag tatttaagat atagcaagga 1320
catggctgat ttttttttat aaaaattggc actaataagt gggtttattg gtcttttcta 1380
attgtataat ttaatttagt acaaagtttg taaaatatca gaggatatat atatattggt 1440
tctacgacat ggtattgcat ttatatcttt ttaattacagt gatctgtgac agcagcagct 1500
tcatgttgta ttttttttac tgaaattgta aaatatccat cttaaagaca tcaactatc 1560
taaaaattgt gtacaggata ttcttttagt ggtggaatta aaatgtacga atacttgctt 1620
tttcaaaaaa atgtattttc tgttaaaagt ttaaagattt ttgctatata ttatggaaga 1680
aaaatgtaat cgtaaatatt aattttgtac ctatattgtg caatacttga aaaaaacggg 1740
ataaaagtat tttgagtcag tgtcttacat gtttaagagg actgaaatag tttatattaa 1800
gtttgtatta aaattcttta aaattaaaaa 1830

```

<210> 190

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005749

<400> 190

```

aaacctctgc tctttctcct aatgccaaagg aattttattt tcctaatatg caggggtcaag 60

```

<210> 191

<211> 1534

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005804

<400> 191

```

ggaagcgcag caactcgtgt ctgagcgccc ggcggaaaac cgaagttgga agtgtctctt 60
agcagcgcgc ggagaagaac ggggagccag catcatggca gaacaggatg tggaaaacga 120
tcttttggat tacgatgaag aggaagagcc ccaggctcct caagagagca caccagctcc 180

```



```

ccctaagaaa gacatcaagg gatcctacgt ttccatccac agctctgggt tccgggactt 240
tctgctgaag cgggagctcc tgcggggccat cgtggactgt ggctttgagc atccttctga 300
gggtccagcat gagtgcattc cccaggccat cctgggcatg gacgtcctgt gccaggccaa 360
gtccgggatg ggcaagacag cgggtcttcgt gctggccacc ctacagcaga ttgagcctgt 420
caacggacag gtgacgggtc tgggtcatgtg ccacacgagg gagctggcct tccagatcag 480
caaggaatat gagcgctttt ccaagtacat gccagcgtc aagggtgtctg tgttcttcgg 540
tgggtctctc atcaagaagg atgaagaagt gttgaagaag aactgtcccc atgtcgtggg 600
ggggaccccc ggccgcaccc tggcgctcgt gcggaatagg agcttcagcc taaagaatgt 660
gaagcacttt gtgctggacg agtgtgacaa gatgctggag cagctggaca tgcggcgagg 720
tgtgcaggag atcttccgcc tgacaccaca cgagaagcag tgcattgatg tcagcgccac 780
cctgagcaag gacatccggc ctgtgtgcag gaagttcatg caggatccca tggagggtgt 840
tgtggacgac gagaccaagc tcacgctgca cggcctgcag cagtactacg tcaaactcaa 900
agacagttag aagaaccgca agctctttga tctcttggat gtgctggagt ttaaccaggt 960
gataatcttc gtcaagtcag tgcagcgtc gatggccctg gccagctcc tcgtggagca 1020
gaacttcccc gccatcgcca tccaccgggg catggcccag gaggagcgcc tgtcacgcta 1080
tcagcagttc aaggatttcc agcggcgagg cctgggtggc accaatctgt ttggccgggg 1140
gatggacatc gagcgagtca acatcgtctt taactacgac atgcctgagg actcggacac 1200
ctacctgcac cgggtggccc gggcggtcgt ctttggcacc aaaggcctag ccatcacttt 1260
tgtgtctgac gagaatgatg ccaaaatcct caatgacgtc caggaccggt ttgaagttaa 1320
tgtggcagaa cttccagagg aaatcgacat ctccacatac atcgagcaga gccggtaacc 1380
accacgtgcc agagccgccc acccgagacc gcccgcatgc agcttcacct cccctttcca 1440
ggcgccactg ttgagaagct agagattgta tgagaataaa cttgttatta tggaaaaaaa 1500
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1534

```

<210> 192
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005804

<400> 192
 gttgagaagc tagagattgt atgagaataa agtggttatta tgaaatgaag aagcctcacc 60

<210> 193
 <211> 1416

<212> DNA
 <213> Homo sapiens

<220>
 <221> Modified_base
 <222> 1 ... 1416
 <223> n = a,c,g, or t

<300>
 <308> NM_005945

<400> 193
 aggaattccg gaattccgga attccgatgg atggaacaga aaataaatct aagtttgggtg 60
 cgaacgccat tctgggggtg tcccttgccg tctgcaaagc tgggtgccgtt gagaaggggg 120
 tcccctgtac cgccacatcg cgtacttggc tggcaacttc gaagtcacatc tgccagtccc 180
 ggcgttcaag tgtcatcatc aatggcggtt ctcatgcttg caacaagctg gccatgcaga 240
 gtctgtcctc ccagtcgggtg cagcaaactc agggaagcca tgccgcattg gagcagagggt 300
 ttaccacaac ctgaagaatg tcatcaagga gaaatatggg aaagatgcca ccaatgtggg 360
 gattttgcgc ggtttgctcc caacatcctg gagaataaag aaggcctgga gctgctgaag 420
 actgctattg gaaagcctgg cctacactgt aaagggtggc atggcatgga cgtagcggcc 480
 tccgagttct tcaggtcagg gaactatgac ctggacttca agtctcccga tgacccagc 540
 aggtacatct cgcctgacca gctggctgac ctgtacaagt ccttcatcaa ggactaccca 600
 gtggtgtcta tcgaagatcc ctttgaccag gatgactggg gagcttcaga agttcacagc 660

```

cagtgcagga atccaggtag tgggggggatg actcacagtg accaacccaa agaggatcgc 720
caagggcgtga acgagaagtc ctgcaactgc ctctctgctca aagtcaacca gattggctcc 780
gtgaccgagt ctcttcaggc gtgcaagctg gcccaggcca atgggtgggg cgtcatggtg 840
tctcatcggt cgggggagac tgaagatacc ttcacgcgtg acctgggtgt ggggctgtgc 900
actggggcag atcaagactg gtgccccttg ccgatcacgc gcttggccaa gtacaaccag 960
ctcctcagaa ttgaagagga gctgggcagc aaggctaagt ttgccggcag gaacttcaga 1020
aacccttgg ccaagtaagc tgtgggcagg caagccttcg gtcacctgtt ggctacagac 1080
ccctcccctg gtgtcagctc aggcagctcg agggcccoga ccaacacttg caggggtccc 1140
tgctagttag cgcccaccgc cgtggagttc gtaccgcttc cttagaactc tacagaagcc 1200
aagctccctg gaagccctgt tggcagctct agctttgcag ttgtgtaatt ggccaagtc 1260
attgtttttc tcgccttact ttccaccaag tgtctagagt catgtgagcc tngtgtcatc 1320
tccgggggtg ccacaggcta gatccccggt gggtttgtgc tcaaaataaa aagcctcagt 1380
gacccatgaa aaaaaaaaaa gaattccgga attccg 1416

```

<210> 194
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005945

<400> 194
 ttgtgtaatt ggccaagtc attgtttttc tcgccttact ttccaccaag tgtctagagt 60

<210> 195
 <211> 961
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006014

<400> 195
 ggcgaccacg gtgtcttcaa aagccccgtc aggggttggt tcttgggggc ggaccgactg 60
 tgggtcagtt tgcaccagcg ctctggaate gatttacgcy cgaaaggcca gaggtttctgg 120
 aggaaaccgc agcctctcaa ccgctgaccg ggtctcagaa ggcccccggc agggcgctt 180
 ggcgggaact gaccacgcgc cagtcaggct ctccagggaac ctgcgcaggc gcgtgtgggc 240
 ggagtcgtgc gcagggggcg gggcttcggg aaggagccac agagagggcg gggcgtagga 300
 cctgcgcttc gggggtggag tcggagcggc gcggcgccgg tcatgcggga cgcggatgca 360
 gacgcaggcg gaggcgctga cggcggggat ggccgggggt gccacagctg ccgccccggc 420
 gtggacacag ccgcagctcc ggccgggtga gctccccag cgcacgcgcc aggtccgggc 480
 agagacgccg cgtctgcggc cagggggtca cgaatgcggc cgcacatatt caccctcagc 540
 gtgcctttcc cgacccctt ggaggcggaa atcgcccatg ggtccctggc accagatgcc 600
 gagccccacc aaagggtggt tgggaaggat ctcacagtga gtggcaggat cctggctcgtc 660
 cgctggaaag ctgaagactg tcgcctgctc cgaatttccg tcatcaactt tcttgaccag 720
 ctttccctgg tggtcggac catgcagcgc tttggggccc ccgtttcccg ctaagcctgg 780
 cctgggcaaa tggagcaggg tcccactttg cgtctccttg taggcagtg gtccatcctt 840
 ccctagggca ggaattccca cagttgctac tttcctggga gggcctcatg ttttatctgg 900
 ttcttaaatg tttgttacta cagaaaataa aactgcgcta ctaaaaaaaa aaaaaaaaaa 960
 a 961

<210> 196
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006014

<400> 196

ggcctcatgt tttatctggt tcttaaagt ttgttactac agaaaataaa actgaggtat 60

<210> 197
 <211> 1648
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006086

<400> 197
 atgcgggaga tctgtcacat ccaggccggc cagtgcggca accagatcgg ggccaagtcc 60
 tgggaagtca tcagtgatga gcatggcatc gaccccgagc gcaactacgt gggcgactcg 120
 gacttgcagc tggagcggat cagcgtctac tacaacgagg cctcttctca caagtacgtg 180
 cctcgagcca ttctgggtgga cctggaaccc ggaacccatgg acagtgtccg ctcagggggc 240
 tttggacatc tcttcaggcc tgacaatttc atctttgggtc agagtggggc cggcaacaac 300
 tgggccaagg gtcactacac ggagggggcg gagctgggtg attcggctct ggatgtgggtg 360
 cggaaggagt gtgaaaactg cgactgcctg cagggtcttc agctgaccca ctcgctgggg 420
 ggggggacgg gctccggcat gggcacgttg ctcatcagca aggtgcgtga ggagtatccc 480
 gaccgcatca tgaacacctt cagcgtcgtg cctcaccaca aggtgtcaga caccgtgggtg 540
 gaaccctaca acgccacgct gtccatccac cagctgggtg aaaacacgga tgaaacctac 600
 tgcacgcaca acgaggcgct ctacgacatc tgcttccgca cctcaagct ggccacggcc 660
 acctacgggg acctcaacca cctgggtatcg gccaccatga gcggagtcac cactccttg 720
 cgcttcccg ggcagctcaa cgctgacctg cgcaagctgg ccgtcaacat ggtgcccttc 780
 ccgcgcctgc acttcttcat gcccggttc gccccctca ccaggcgggg cagccagcag 840
 taccggggcc tgaccgtgcc cgagctcacc cagcagatgt tcgatgccaa gaacatgatg 900
 gccgcctgcg acccgcgcca cggccgctac ctgacgggtg ccaccgtgtt ccggggccgc 960
 atgtccatga aggaggtgga cgagcagatg ctggccatcc agagcaagaa cagcagctac 1020
 ttcgtggagt ggatcccaa caacgtgaag gtggccgtgt gtgacatccc gcccgcggc 1080
 ctcaagatgt cctccacctt catcggaac agcacggcca tccaggagct gttcaagcgc 1140
 atctccgagc agttcacggc catgttccgg cgcaaggcct tcttgcactg gtacacgggc 1200
 gagggcattg acgagatgga gtaccacgag gccgagagca acatgaacga cctggtgtcc 1260
 gaggaccagc agtaccagga cgccacggcc gaggaagagg gcgagatgta cgaagacgac 1320
 gaggaggagt cggaggccca gggccccaag tgaaactgct cgcagctgga gtgagaggca 1380
 ggtggcgggc gggggccgaag ccagcagtgct ctaaaccccc ggagccatct tgctgccgac 1440
 accctgcttt ccccatcgcc ctagggtctc cttgccgccc tctgcagta tttatggcct 1500
 cgtcctcccc cacctaggcc acgtgtgagc tgctcctgtc tctgtcttat tgcagctcca 1560
 ggctgtacgt tttacggttt tgttttttac tggtttgtgt ttatatatttc ggggatactt 1620
 aataaatcta ttgtgtcag ataccctt 1648

<210> 198
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006086

<400> 198
 tttttactgg tttgtgttta ttttttcggg gatacttaat aaatctattg ctgtcagata 60

<210> 199
 <211> 3074
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006096

<400> 199
 aacaaacctc gctgggtcc cagctgggtg tgaagctcgt cagttcacca tccgccctcg 60

```

gcttccgcggg ggcgcctggggc cgccagccctc ggcaccgtcc tttcctttct ccctcgcgtt 120
aggcagggtga cagcaggggac atgtctcggg agatgcagga tgtagacctc gctgagggtga 180
agccttttggg ggagaaaaggg gagaccatca ccggcctcct gcaagagttt gatgtccagg 240
agcaggacat cgagacttta catggctctg ttcacgtcac gctgtgtggg actcccaagg 300

gaaaccgggcc tgtcatcctc acctaccatg acatcggcac gaaccacaaa acctgctaca 360
acccccctctt caactacgag gacatgcagg agatcaccca gcactttgcc gtctgccacg 420
tggaacgcccc tggccagcag gacgggcgag cctccttccc cgcagggtac atgtaccctt 480
ccatggatca gctggctgaa atgcttcctg gagtccttca acagtttggg ctgaaaagca 540
ttattggcat gggaacagga gcaggcgctt acatcctaac tcgatttgct ctaaaacaacc 600
ctgagatggt ggagggcctt gtccttatca acgtgaaccc ttgtgcgga ggctggatgg 660
actgggcgcg ctccaagatc tcaggatgga cccaagctct gccggacatg gtggtgtccc 720
accttttttgg gaaggaagaa atgcagagta acgtggaagt ggtccacacc taccgccagc 780
acattgtgaa tgacatgaac cccggcaacc tgcacctgtt catcaatgcc tacaacagcc 840
ggcgcgacct ggagattgag cgaccaatgc cgggaaccca cacagtcacc ctgcagtgcc 900
ctgctctgtt ggtggttggg gacagctcgc ctgcagtgga tgccgtggtg gagtgcactt 960
caaaattgga cccaacaaag accactctcc tcaagatggc ggactgtggc ggcctcccgc 1020
agatctccca gccggccaag ctgcctgagg ccttcaagta cttcgtgcag ggcattggat 1080
acatgccctc gcttagcatg accgcctga tgggtcccg cacagcctct ggttccagcg 1140
tcacttctct ggatggcacc cgcagccgct cccacaccag cgagggcacc cgaagccgct 1200
cccacaccag cgagggcacc cgcagccgct cgcacaccag cgagggggcc cacctggaca 1260
tcaccccaaa ctcggtgct gctgggaaca gcgcggggc caagtccatg gaggtctcct 1320
gctaggcggc ctgcccagct gccgcccccg gactctgac tctgtagtgg cccctcctc 1380
cccgccctt tttcgcccc tgctgcca actgcgcta actcggattt aatccaaagc 1440
ttattttgta agagtgcgt ctggtggaga caaatgaggt ctattacgtg ggtgccctct 1500
ccaaaggcgg ggtggcgggt gaccaaagga aggaagcaag catctccgca tgcctcctc 1560
ttccattaac cagtggccgg ttgccactct cctccccctc ctccagagaca ccaaactgcc 1620
aaaaacaaga cgcgtagcag cacacacttc acaaagccaa gcctaggccg ccttgagcat 1680
cctggttcaa acgggtgcct ggtcagaagg ccagccgccc acttcccgtt tctcttttaa 1740
ctgaggagaa gctgatccag tttccggaaa caaaatcctt ttctcatttg gggagggggg 1800
taatagtgc atgcaggcac ctctttttaa caggcaaaac aggaaggggg aaaaggtggg 1860
attcatgtcg aggttagagg catttggaa aacaaatcta cgtagttaac ttgaagaaac 1920
cgatttttaa agttggtgca tctagaaagc tttgaatgca gaagcaaaac agcttgattt 1980
ttctagcatc ctcttaatgt gcagcaaaag caggcgacaa aatctcctgg ctttacagac 2040
aaaaatattt cagcaaacgt tgggcatcat ggtttttgaa ggcttttagt ctgctttctg 2100
cctctcctcc acagcccaaa cctcccaccc ctgatacatg agccagtgat tattcttggt 2160
cagggagaag atcatttaga tttgttttgc attccttaga atggagggca acattccaca 2220
gctgccctgg ctgtgatgag tgtccttgca ggggcgggag taggagcact ggggtggggg 2280
tggaattggg gttactcgat gtaagggatt ccttggtgtt gtgttgagat ccagtgcagt 2340
tgtgatttct gtggatccca gcttggttcc aggaattttg tgtgattggc ttaaattccag 2400
ttttcaatct tcgacagctg ggctggaacg tgaactcagt agctgaacct gtctgacccg 2460
gtcacgttct tggatcctca gaactctttg ctcttgctcg ggtgggggtg ggaactcacg 2520
tggggagcgg tggctgagaa aatgtaagga ttctggaata catattccat gggactttcc 2580
ttccctctcc tgcttcctct tttcctgctc cctaaccctt cgccgaatgg ggcagcacca 2640
ctgacgtttc tgggcggcca gtgcggctgc caggttcctg tactactgcc ttgtactttt 2700
cattttggct caccgtggat tttctcatag gaagtttggc cagagtgaat tgaatattgt 2760
aagtcagcca ctgggacccg aggatctctg ggaccccgca gttgggagga ggaagtagtc 2820
cagccttcca ggtggcgtga gaggcaatga ctcgttacct gccgcccatc accttgagg 2880
ccttccctgg ccttgagtag aaaagtcggg gatcggggca agagaggctg agtacggatg 2940
ggaaactatt gtgcacaagt ctttcagag gagtttctta atgagatatt tgtatttatt 3000
tccagaccaa taaatttgta actttgcagc ggaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3060
aaaaaaaaaa aaaa 3074

```

<210> 200

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006096

<400> 200
gagtacggat gggaaactat tgtgcacaag tctttccaga ggagtttctt aatgagatat 60

<210> 201
<211> 2148
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006115

<400> 201
gcttcagggt acagctcccc cgcagccaga agccgggctt gcagcgcctc agcaccgctc 60
cgggacaccc caccgccttc ccaggcgtga cctgtcaaca gcaacttcgc ggtgtggtga 120
actctctgag gaaaaacat tttgattatt actctcagac gtgcgtggca acaagtgact 180
gagacctaga aatccaagcg ttggaggctc tgaggccagc ctaagtcgct tcaaaatgga 240
acgaaggcgt ttgtggggtt ccattcagag ccgatacatc agcatgagtg tgtggacaag 300
cccacggaga cttgtggagc tggcagggca gagcctgctg aaggatgagg cctggccat 360
tgccgccctg gagttgctgc ccagggagct cttcccgcca ctcttcattg cagcctttga 420
cgggagacac agccagaccc tgaaggcaat ggtgcaggcc tggcccttca cctgcctccc 480
tctgggagtg ctgatgaagg gacaacatct tcacctggag accttcaaag ctgtgcttga 540
tggacttgat gtgctccttg ccaggagggt tcgccccagg aggtggaaac ttcaagtgtc 600
ggatttacgg aagaactctc atcaggactt ctggactgta tggcttgga acagggccag 660
tctgtactca tttccagagc cagaagcagc tcagcccatg acaaagaagc gaaaagtaga 720
tggttttgag acagaggcag agcagccctt cattccagta gaggtgctcg tagacctgtt 780
cctcaaggaa ggtgcctgtg atgaattgtt ctctacctc attgagaaag tgaagcgaag 840
gaaaaatgta ctacgcctgt gctgtaagaa gctgaagatt tttgcaatgc ccatgcagga 900
tatcaagatg atcctgaaaa tgggtgcagct ggactctatt gaagatttgg aagtgacttg 960
tacctggaag ctaccacact tggcgaaatt ttctccttac ctgggccaga tgattaatct 1020
gcgtagactc ctctctccc acatccatgc atcttcctac atttccccgg agaaggaaga 1080
gcagtataac gccagttca cctctcagtt cctcagctcg cagtgcctgc aggtctctta 1140
tgtggactct ttatTTTTCC ttagaggccg cctggatcag ttgctcaggc acgtgatgaa 1200
ccccttgga accctctcaa taactaactg ccggctttcg gaaggggatg tgatgcatct 1260
gtcccagagt cccagcgtca gtcagctaag tgtcctgagt ctaagtgggg tcatgctgac 1320
cgatgtaagt cccgagcccc tccaagctct gctggagaga gcctctgcca ccctccagga 1380
cctggtcttt gatgagtgtg ggaatcacgga tgatcagctc cttgccctcc tgccttcctt 1440
gagccactgc tcccagctta caaoccttaag cttctacggg aattccatct ccatactctc 1500
cttgacagat ctcttcgagc acctcatcgg gctgagcaat ctgaccacag tgctgtatcc 1560
tgtccccctg gagagttatg aggacatcca tggtaacctc cacctggaga ggcttgccca 1620
tctgcatgcc aggtcaggg agttgctgtg tgagttgggg cggcccagca tggctctggct 1680
tagtgccaac cctgtcctc actgtgggga cagaaccttc tatgaccggg agccatcct 1740
gtgccctgt ttcatgccta actagctggg tgcacatata aaatgcttca ttctgcatac 1800
ttggacacta aagccaggat gtgcatgcat cttgaagcaa caaagcagcc acagtttcag 1860
acaaatgttc agtggtgagt aggaaaacat gttcagttag gaaaaaacat tcagacaaat 1920
gttcagttag gaaaaaaagg ggaagttggg gataggcaga tgttgacttg aggagttaat 1980
gtgatctttg gggagataca tcttatagag ttagaaatag aatctgaatt tctaaaggga 2040
gattctggct tgggaagtac atgtaggagt taatccctgt gtagactgtt gtaaagaaac 2100
tgttgaaaat aaagagaagc aatgtgaagc aaaaaaaaaa aaaaaaaa 2148

<210> 202
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006115

<400> 202
tggggagata catcttatag agttagaaat agaactctgaa tttctaaagg gagattctgg 60

<210> 203
 <211> 1051
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006332

<400> 203
 ggaccgcgcg ctggttaaag gcgcttattt cccaggcagc cgctgcagtc gccacacctt 60
 tgcccctgct gcgatgaccc tgtcgccact tctgctgttc ctgccaccgc tgctgctgct 120
 gctggacgtc cccacggcgg cggtgcaggc gtcccctctg caagcgtagg acttccttgg 180
 gaatgggcca ccagttaact acaagacagg caatctatac ctgcgggggc ccctgaagaa 240
 gtccaatgca ccgcttgtca atgtgacctt ctactatgaa gactgtgctg gtggctgccg 300
 agccttccctg atccgggagc tcttcccaac atggctgttg gtcatggaga tcctcaatgt 360
 cagcttggtg ccctacggaa acgcacagga acaaaatgtc agtggcaggt gggagttcaa 420
 gtgccagcat ggagaagagg agtgcaaatt caacaagggt gaggcctgcg tgttgatga 480
 acttgacatg gagctagcct tcctgaccat tgtctgcatg gaagagtttg aggacatgga 540
 gagaagtctg ccactatgcc tgcagctcta cgcccagggt ctgtcgccag acactatcat 600
 ggagtgtgca atgggggacc gcggcatgca gctcatgcac gccaacgccc agcggacaga 660
 tgctctccag ccaccacacg agtatgtgcc ctgggtcacc gtcaatggga aacccttggg 720
 agatcagacc cagctcctta cccttgtctg ccagttgtac cagggcaaga agccggatgt 780
 ctgcccttcc tcaaccagct ccctcaggag tgtttgcttc aagtgtatgc cggtgagctg 840
 cggagagctc atggaaggcg agtgggaacc cggctgcctg cctttttttc tgatccagac 900
 cctcggcacc tgctacttac caactggaaa attttatgca tcccatgaag ccagataca 960
 caaaattcca ccccatgatc aagaatcctg ctccactaag aatggtgcta aagtaaaact 1020
 agtttaataa gcaaaaaaaaa aaaaaaaaaa a 1051

<210> 204
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006332

<400> 204
 aaattccacc cctagatcaa gaatcctgct ccactaagaa tgggtgctaaa gtaaaactag 60

<210> 205
 <211> 1714
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006417

<400> 205
 ggggcatttt gtgcctgcct agctatccag acagagcagc taccctcagc tctagctgat 60
 actacagaca gtacaacaga tcaagaagta tggcagtgac aactcgtttg acacggttgc 120
 acgaaaagat cctgcaaaat cattttggag ggaagcggct tagccttctc tataagggtg 180
 gtgtccatgg attccgtaat ggagttttgc ttgacagatg ttgtaatcaa gggcctactc 240
 taacagtgat ttatagtga gatcatatta ttggagcata tgcggaagag agttaccagg 300
 aaggaaagta tgcttccatc atcctttttg cacttcaaga tactaaaatt tcagaatgga 360
 aactaggact atgtacacca gaaacactgt tttgttgtga tgttacaaaa tataactccc 420
 caactaattt ccagatagat ggaagaaata gaaaagtgat tatggactta aagacaatgg 480
 aaaatccttg acttgctcaa aattgtacta tctctattca ggattatgaa gtttttcgat 540
 gcgaagattc actggatgaa agaaagataa aaggggtcat tgagctcagg aagagcttac 600
 tgtctgcctt gagaacttat gaaccatatt gatccctggg tcaacaaata cgaattctgc 660
 tgctgggtcc aattggagct gggaagtcca gctttttcaa ctcagtgagg tctgttttcc 720

```

aagggcatgt aacgcatcag gctttggtgg gcactaatac aactgggata tctgagaagt 780
ataggacata ctctattaga gacgggaaag atggcaaata cctgccgttt attctgtgtg 840
actcactggg gctgagttag aaagaaggcg gcctgtgcag ggatgacata ttctatatct 900
tgaacggtaa cattcgtgat agataccagt ttaatcccat ggaatcaatc aaattaaatc 960
atcatgacta cattgattcc ccatcgctga aggacagaat tcattgtgtg gcatttgtat 1020
ttgatgccag ctctattcaa tactttctct ctcagatgat agtaaagatc aaaagaattc 1080
gaagggagtt ggtaaacgct ggtgtggtac atgtggcctt gctcactcat gtggatagca 1140
tggatttgat tacaaaaggt gaccttatag aaatagagag atgtgagcct gtgaggtcca 1200
agctagagga agtccaaaga aaacttggat ttgctctttc tgacatctcg gtggtagca 1260
attattctctc tgagtgggag ctggaccctg taaaggatgt tctaattctt tctgctctga 1320
gacgaatgct atgggctgca gatgacttct tagaggattt gccttttgag caaataggga 1380
atctaaggga ggaattatc aactgtgcac aaggaaaaaa atagatatgt gaaagggtca 1440
cgtaaatttc ctcacatcac agaagattaa aattcagaaa ggagaaaaca cagaccaaag 1500
agaagtatct aagaccaaag ggatgtgttt tattaatgtc taggatgaag aaatgcatag 1560
aacattgtag tacttgtaaa taactagaaa taacatgatt tagtcataat tgtgaaaaat 1620
agtaataatt tttcttggtt ttatgttctg tatctgtgaa aaaataaatt tcttataaaa 1680
ctcggaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1714

```

<210> 206

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006417

<400> 206

```

atgacatatt ctatatcttg aacggtaaca ttcgtgatag ataccagttt aatcccatgg 60

```

<210> 207

<211> 3791

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006461

<400> 207

```

acagacggcg ggtgaacatg gcgtcctcga cttggtctga gacgtgatag gcctgccttc 60
tggttgaaga tgtggcgagt gaaaaaactg agcctcagcc tgtgcgcttc gcccagacg 120
ggaaaacccat ctatgagaac tcctctccgt gaacttaccc tgcagcccg tgccctcacc 180
acctctggaa aaagatcccc cgttctctcc tcgtgacccc catcactgtg caagctgggg 240
ctgcaggaag gcagcaacaa ctogtctcca gtggattttg taaataacaa gaggacagac 300
ttatcttcag aacatttcag tcattcctca aagtggctag aaacttgtca gcatgaatca 360
gatgagcagc ctctagatcc aattccccc attagctcta ctcctaaaac gtctgaggaa 420
gcagtagacc cactgggcaa ttatatggtt aaaaccatcg tccttgtagc atctccactg 480
gggcagcaac aagacatgat atttgaggcc cgttttagata ccattggcaga gacaaacagc 540
atatctttta atggaccttt gagaacagac gatctggtga gagaggaggt ggcacctgc 600
atgggagaca ggttttcaga agttgtctgt gtatctgaga aacctatctt tcaggaatct 660
ccgtcccatc tcttagagga gtctccacca aatccctgtt ctgaacaact acattgctcc 720
aaggaaagcc tgagcagtag aactgaggct gtgcgtgagg acttagtacc ttctgaaagt 780
aacgccttct tgccctctctc tgttctcttg ctttccctct caactgcctt ggcagcagat 840
ttccgtgtca atcatgtgga ccagaggag gaaattgtag agcatggagc tatggaggaa 900
agagaaatga ggtttccac acatcctaag gagtctgaaa cagaagatca agcactgtc 960
tcaagtgtgg aagatattct gtccacatgc ctgacaccaa atctagtaga aatggaatcc 1020
caagaagctc caggcccagc agtagaagat gttggttaga ttcttggtc tgatacagag 1080
tcttggtatg cccactggc ctggctggaa aaagggtgaa atacctcgt catgctggaa 1140
aatctccgcc aaagcttata ccttccctcg atgcttcggg atgctgcaat tggcactacc 1200
cctttctcta cttgtcggg ggggacttgg tttactcctt cagcaccaca ggaaaagagt 1260
acaaacacat cccagacagg cctggttggc accaagcaca gtacttctga gacagagcag 1320
ctcctgtgtg gccggcctcc agatctgact gccttgtctc gacatgactt ggaagataac 1380

```

```

ctgctgagct ctcttgtcat tgtggagttt ctctcccgcc agcttcggga ctggaagagc 1440
cagctggctg tccctcaccg agaaacccag gacagtagca cacagactga cacatctcac 1500
agtgggataa ctaataaaact tcagcatcctt aaggagagcc atgagatggg acaggcccta 1560
cagcaggcca gaaatgtcat gcaatcatgg gtgcttatct ctaaagagct gatatccttg 1620
cttcacctat ccctgttgca tttagaagaa gataagacta ctgtgaatca ggagtctcgg 1680
cgtgcagaaa cattggtctg ttgctgtttt gatttgctga agaaattgag ggcaaagctc 1740
cagagcctca aagcagaaaag ggaggaggca aggcacagag aggaaatggc tctcagaggc 1800
aaggatgctg cagagatagt gttggaggct ttctgtgcac acgccagcca gcgcctcagc 1860
cagctggaac aggacctagc atccatgcgg gaattcagag gccttctgaa ggatgccag 1920
acccaactgg tagggcttca tgccaagcaa gaagagctgg ttcagcagac agtgagtctt 1980
acttctacat tgcaacaaga ctggaggctc atgcaactgg attatacaac atggacagct 2040
ttgctgagtc ggtcccgcac actcacagag aaactcacag tcaagagcca gcaagccctg 2100
caggaacgtg atgtggcaat tgaggaaaag caggagggtt ctagggtgct ggaacaagtc 2160
tctgcccagt tagaggagtg caaaggccaa acagaacaac tggagttgga aacatttcgt 2220
ctagcaacag atctccgggc tcagttgcag attctggcca acatggacag ccagctaaaa 2280
gagctacaga gtcagcatat ccattgtgcc caggacctgg ctatgaagga tgagttactc 2340
tgccagctta cccagagcaa tgaggagcag gctgctcaat gcgtaaagga agagatggca 2400
ctaaaacaca agcaggcaga actgcagcag caacaagctg ccctggccaa agagggtgcg 2460
gacctgaaag agaccttga gtttgcagac caggagaatc aggttgctca cctggagctg 2520
ggtcagggtt agtgtcaatt gaaaaccaca ctggaagtgc tccgggagcg cagcttgcag 2580
tgtgagaacc tcaaggacac tgtagagaa ctaacggcta aactggccag caccatagca 2640
gataaccagg agcaagatct ggagaaaaca cggcagtag ctcaaaagct agggctgctg 2700
actgagcaac tacagagcct gactctcttt ctacagacaa aactaaagga gaagactgaa 2760
caagagacct ttctgtctgag tacagcctgt cctcccaccc aggaacaccc tctgcctaact 2820
gacaggacct tcctgggaag catcttgaca gcagtggcag atgaagagcc agaatacaact 2880
cctgtgccct tgcttggaa gacaaagagt gctttcaccc gagtagcatc aatgggtttcc 2940
cttcagcccg cagagacccc aggcattggg gagagcctgg cagaaatgag tattatgact 3000
actgagcttc agagtctttg ttccctgcta caagagtcta aagaagaagc catcaggact 3060
ctgcagcgaa aaatttgtga gctgcaagct aggctgcagg cccaggaaga acagcatcag 3120
gaagtccaga aggcaaaaaga agcagacata gagaagctga accaggcctt gtgcttgcgc 3180
tacaagaatg aaaaggagct ccagggaagt atacagcaga atgagaagat cctagaacag 3240
atagacaaga gtggcgagct cataagcctt agagaggagg tgaccacact taccgctca 3300
cttcggcgct cgagacaga gaccaaagt ctccaggagg cctggcaggc cagctggact 3360
ccaactgcca gcctatggc accaattgga tccaggagaa agtgtggctc tctcaggagg 3420
tggaacaaact gagagtgatg ttccctggaga tgaaaaatga gaaggaaaac tcctgatcaa 3480
gttcagagc ccatagaaat atcctagagg agaaccctcg gcgctctgac aaggagttag 3540
aaaaactaga tgacattgtt cagcatattt ataagacct gctctctatt ccagagggtg 3600
tgaggggatg caaagaacta cagggattgc tggaatttct gagctaagaa actgaaagcc 3660
agaatttgtt tcacctcttt ttacctgcaa taccacctta cccaataacc aagaccaact 3720
ggcatagagc caactgagat aaatgctatt taaataaagt gtatttaatg aaaaaaaaaa 3780
aaaaaaaaaa a 3791

```

<210> 208
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006461

<400> 208
 ctgacaagga gttagaaaaa ctagatgaca ttgttcagca tatttataag accctgctct 60

<210> 209
 <211> 2856
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006516

<400> 209

```

tagtcgcggg tccccgagtg agcacgccag ggagcaggag accaaacgac gggggctcgga 60
gtcagagtcg cagtgggagtg ccccggaaccg gagcacgagc ctgagcggga gagcgccgt 120
cgcacgcccg tcgccacccg cgtacccggc gcagccagag ccaccagcgc agcgctgcca 180
tggagcccag cagcaagaag ctgacgggtc gcctcatgct ggctgtggga ggagcagtg 240
ttggctccct gcagtttggc tacaacactg gagtcatcaa tgccccccag aagggtgatcg 300
aggagttcta caaccagaca tgggtccacc gctatgggga gagcatcctg cccaccacgc 360
tcaccacgct ctggtccctc tcagtggcca tcttttctgt tgggggcatg attgggtcct 420
tctctgtggg ccttttcgtt aaccgctttg gccggcgga ttcaatgctg atgatgaacc 480
tgctggcctt cgtgtccgcc gtgctcatgg gcttctcgaa actgggcaag tcttttgaga 540
tgctgatact gggccgcttc atcatcggtg tgtactgcgg cctgaccaca ggcttcgtgc 600
ccatgtatgt ggggtgaagtg tcaccacag cctttcgtgg ggccctgggc accctgcacc 660
agctgggcat cgtcgtcggc atcctcatcg ccaggtgtt cggcctggac tccatcatgg 720
gcaacaagga cctgtggccc ctgctgctga gcatcatctt catcccgcc ctgctgcagt 780
gcatcgtgct gcccttctgc cccgagagtc cccgcttctt gctcatcaac cgcaacgagg 840
agaaccgggc caagagtgtg ctaaagaagc tgccggggac agctgacgtg acccatgacc 900
tgcaggagat gaaggaagag agtcggcaga tgatgcggga gaagaaggct accatcctgg 960
agctgttccg ctcgccgcc taccgccagc ccactctcat cgtgtggtg ctgcagctgt 1020
cccagcagct gtctggcatc aacgctgtct tctattactc cacgagcatc ttcgagaagg 1080
cgggggtgca gcagcctgtg tatgccacca ttggctccgg tatcgtcaac acggccttca 1140
ctgtcgtgtc gctgtttgtg gtggagcgag caggccggcg gacctgcac ctcatagggc 1200
tcgttggcat ggcgggttgt gccatactca tgaccatgc gctagcactg ctggagcagc 1260
taccctggat gtcctatctg agcatcgtgg ccactcttgg ctttgtggcc tcttttgaag 1320
tgggtccttg ccccatccca tggttcatcg tggctgaact cttcagccag ggtccacgtc 1380
cagctgccat tgcggttgca ggcttctcca actggacctc aaatttcatt gtgggcatgt 1440
gcttccagta tgtggagcaa ctgtgtggtc cctacgtctt catcatcttc actgtgctcc 1500
tggttctgtt cttcatcttc acctacttca aagttcctga gactaaaggc cggaccttcg 1560
atgagatcgc ttccggcttc cggcaggggg gagccagcca aagtgataag acaccgagg 1620
agctgttcca tccctgggg gctgattccc aagtgtagt cgcccagat caccagccc 1680
gcctgtctcc agcagcccta aggatctctc aggagcacag gcagctggat gagacttcca 1740
aacctgacag atgtcagccg agccgggctt ggggctcctt tctccagcca gcaatgatgt 1800
ccagaagaat attcaggact taacggctcc aggtttttat aattttttta ttactgattt tgttattttt 1860
aaatctattc agacaagcaa caggttttat cctgaatgg ttccatgctt 1920
atatcagcct gagtctcctg tgcccacatc ccaggcttca cctgaatgg ttccatgctt 1980
gaggttgag actaagccct gtccagacac ttgcttctt caccagcta atctgtagg 2040
ctggacctat gtcctaagga cacactaatc gaactatgaa ctacaaagct tctatcccag 2100
gaggtggcta tggccacccg ttctgtgtgg ctggatctcc ccactctagg ggtcaggctc 2160
cattaggatt tgcccccttc catctcttcc taccacaacca ctcaaattaa tctttcttta 2220
cctgagacca gttgggagca ctggagtga gggaggagag gggaagggcc agtctgggct 2280
gccgggttct agtctccttt gcactgaggg ccacactatt accatgagaa gagggcctgt 2340
gggagcctgc aaactcactg ctcaagaaga catggagact cctgcctgt tgtgtataga 2400
tgcaagatat ttatatatat ttttggttgt caatattaaa tacagacact aagttatagt 2460
atatctggac aagccaactt gtaaatatcac cacctcactc ctgttactta cctaaacaga 2520
tataaatggc tggtttttag aaacatgggt ttgaaatgct tgtggattga gggtaggagg 2580
tttggtggg agtgagacag aagtaagtgg ggttgcaacc actgcaacgg cttagacttc 2640
gactcaggat ccagtccttc acacgtacct ctcatcagtg tctcttgtt caaaaatctg 2700
tttgatccct gttaccaga gaatatatac attctttatc ttgacattca aggcatttct 2760
atcacatatt tgatagttgg tgttcaaaaa aacactagtt ttgtgccagc cgtgatgctc 2820
aggcttgaaa tcgcattatt ttgaatgtga agggaa 2856

```

<210> 210

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006516

<400> 210

```

aaacagatat aaatggctgg ttttttagaaa catggttttg aaatgcttgt ggattgaggg 60

```

<210> 211
 <211> 576
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006607

<400> 211
 atggctactc tgatctacgt tgataaggaa attgggagaac caggcaccgc tgtggctgcc 60
 aaggatgtgc tgaagctgga gtctagacct tcaatcaaag cattagatgg gatatctcaa 120
 gttttaacac cacgttttgg caaaacatac gatgtcccat cagccttacc taaagctacc 180
 agaaaggctt tgggcactgt caacagagct acagaaaagt cagtaaagac caatggaccc 240
 agaaaacaaa aacagccaag cttttctgcc aaaaagatga ccgagaagac tgttaaaaca 300
 aaaagtctctg ttctgcctc agatgacgcc tatccagaaa tagaaaaatt ctttcccttc 360
 aatcttctag actttgagag ttttgacctg cctgaagagc gccagattgc acacctcccc 420
 ttgagtggag tgcctctcat gatccttgat gaggagggag agcttgaaaa gctgtttcag 480
 ctgggcccc cttcacctgt gaaaatgccc tctccaccat gggaatgcaa tctgtttgca 540
 gtctccttca agcattctgt cgaccctgga tgttga 576

<210> 212
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006607

<400> 212
 cgcctatcca gaaatagaaa aattctttcc cttcaatctt ctagactttg agagttttga 60

<210> 213
 <211> 2058
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006820

<400> 213
 gcacgaggaa gccacagatc tcttaagaac tttctgtctc caaacctgtg ctgctcgata 60
 aatcagacag aacagttaat cctcaattta agcctgatct aacccttaga aacagatata 120
 gaacaatgga agtgacaaca agattgacat ggaatgatga aaatcatctg cgcaactgct 180
 tggaaatgtt tctttgagtc ttctctataa gtctagtgtt catggaggta gcattgaaga 240
 tatggttgaa agatgcagcc gtcagggatg tactataaca atggcttaca ttgattacaa 300
 tatgattgta gcctttatgc ttggaaatta tattaattta cgtgaaagtt ctacagagcc 360
 aaatgattcc ctatggtttt cacttcaaaa gaaaaatgac accactgaaa tagaaacttt 420
 actcttaaat acagcaccaa aaattattga tgagcaactg gtgtgtcgtt tatcgaaaac 480
 ggataatttc attatatgtc gagataataa aattttatcta gataaaatga taacaagaaa 540
 cttgaaacta aggtttttatg gccaccgtca gtattttggaa tgtgaagttt ttcgagttga 600
 aggaattaag gataacctag acgacataaa gaggataaatt aaagccagag agcacagaaa 660
 taggcttcta gcagacatca gagactatag gccctatgca gacttgggtt cagaaattcg 720
 tattcttttg gtgggtccag ttgggtctgt aaagtccagt tttttcaatt cagtcaagtc 780
 tatttttcat ggccatgtga ctggccaagc cgtagtgggg tctgatacca ccagcataac 840
 cgagcgggtat aggatattat ctgttaaaga tggaaaaaat ggaaaatctc tgccatttat 900
 gttgtgtgac actatggggc tagatggggc agaaggagca ggactgtgca tggatgacat 960
 tccccacatc ttaaaagggt gtatgccaga cagatatcag tttaattccc gtaaaccaat 1020
 tacacctgag cattctactt ttatcacctc tccatctctg aaggacagga ttactgtgt 1080
 ggcttatgtc ttagacatca actctattga caatctctac tctaaaatgt tggcaaaagt 1140
 gaagcaagtt cacaagaag tattaaactg tggatatgca tatgtggcct tgcttactaa 1200

```

agtggatgat tgcagtgagg ttcttcaaga caacttttta aacatgagta gatctatgac 1260
ttctcaaagc cgggtcatga atgtccataa aatgctaggc attcctatctt ccaatatttt 1320
gatggttggg aattatgctt cagatttgga actggacccc atgaaggata ttctcatcct 1380
ctctgcactg aggcagatgc tgcgggctgc agatgatttt ttagaagatt tgcctcttga 1440
ggaaaactggt gcaattgaga gagcggttaca gccctgcatt tgagataagt tgccttgatt 1500
ctgacatttg gccagcctg tactggtgtg ccgcaatgag agtcaatctc tattgacagc 1560
ctgcttcaga ttttgctttt gtctgttttg ccttctgtcc ttggaacagt catactctaa 1620
gttcaaagc caaaacctga gaagcgggtg gctaagatag gtctactgc aaaccacccc 1680
tccatatttc cgtaccattt acaattcagt ttctgtgaca tcttttttaa cactggagg 1740
aaaaatgaga tattctctaa tttattcttc tataacactc tatatagagc tatgtgagta 1800
ctaatacacat tgaataatag ttataaaatt attgtataga catctgcttc ttaaacagat 1860
tgtgagttct ttgagaaaca gcgtggattt tacttatctg tgtattcaca gagcttagca 1920
cagtgcctgg taatgagcaa gcatacttgc cactactttt ccttcccact ctctccaaca 1980
tcacattcac tttaaatttt tctgtatata gaaaggaaaa ctagcctggg caacatgatg 2040
aaaccccatc tccactgc 2058

```

<210> 214

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006820

<400> 214

```

tgagttcttt gagaaacagc gtggatttta cttatctgtg tattcacaga gcttagcaca 60

```

<210> 215

<211> 2825

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006845

<400> 215

```

gcgaaattga ggtttcttgg tattgcgcgt ttctcttctt tgctgactct ccgaatggcc 60
atggactcgt cgcttcaggc ccgcctgttt cccggctctg ctatcaagat ccaacgcagt 120
aatggtttta ttacagtgcc caatgtaagg actgtgaact tggagaaatc ctgtgtttca 180
gtggaatggg cagaaggagg tgccacaaag ggcaaaagaga ttgattttga tgatgtggct 240
gcaataaacc cagaactctt acagcttctt cctttacatc cgaaggacaa tctgcccttg 300
caggaaaatg taacaatcca gaaacaaaaa cggagatccg tcaactccaa aattcctgct 360
ccaaaagaaa gtcttcgaag ccgtccact cgcattgtcca ctgtctcaga gcttcgcatt 420
acggctcagg agaatgacat ggaggtggag ctgcctgcag ctgcaaaact ccgcaagcag 480
ttttcagttc ctctgcccc cactaggcct tccgtgcccgt cagtggctga aataccattg 540
aggatggtca gcgaggagat ggaagagcaa gtccattcca tccgtggcag ctcttctgca 600
aaccctgtga actcagttcg gaggaatca tgtcttgtga aggaagtggg aaaaatgaag 660
aacaagcgag aagagaagaa ggcccagaac tctgaaatga gaatgaagag agctcaggag 720
tatgacagta gttttccaaa ctgggaattt gcccgaaatg ttaaagaatt tggggtact 780
ttggaatgtc atccacttac tatgactgat cctatcgaag agcacagaat atgtgtctgt 840
gttaggaaac gccactgaa taagcaagaa ttggccaaga aagaaattga tgtgatttcc 900
attcctagca agtgtctcct cttggtacat gaacccaagt tgaaagtggg cttacaagaag 960
tatctggaga accaagcatt ctgctttgac tttgcatttg atgaaacagc ttccaatgaa 1020
gttgtctaca ggttcacagc aaggccactg gtacagacaa tctttgaagg tggaaaagca 1080
acttgttttg catatggcca gacaggaagt ggcaagacac atactatggg cggagacctc 1140
tctgggaaag ccagaaatgc atccaaaggg atctatgcc tggcctcccg ggacgtcttc 1200
ctctgaaga atcaaccctg ctaccggaag ttgggcctgg aagtctatgt gacattcttc 1260
gagatctaca atgggaagct gtttgacctg ctcaacaaga aggccaagct gcgcgtgctg 1320
gaggacggca agcaacaggt gcaagtgggt gggctgcagg agcatctggt taactctgct 1380
gatgatgtca tcaagatgct cgacatgggc agcgcctgca gaacctctgg gcagacattt 1440
gccaaactcca attcctcccg ctcccacgag tgcttccaaa ttattcttct agctaaaggg 1500

```



```

agaatgcatg gcaagttctc tttggttagat ctggcagggg atgagcggagg cgcagacact 1560
tccagtgcctg accggcagac ccgcatggag ggcgcagaaa tcaacaagag tctcttagcc 1620
ctgaaggagt gcatcagggc cctgggacag aacaaggctc acaccccggt ccgtgagagc 1680
aagctgacac aggtgctgag ggactccttc attggggaga actctaggac ttgcatgatt 1740
gccacgatct caccaggcat aagctcctgt gaataactt taaacaccct gagatatgca 1800
gacagggcca aggagctgag cccccacagt gggcccagtg gagagcagtt gattcaaatg 1860
gaaacagaag agatggaagc ctgctctaac ggggcgctga ttccaggcaa tttatccaag 1920
gaagaggagg aactgtcttc ccagatgtcc agctttaacg aagccatgac tcagatcagg 1980
gagctggagg agaaggctat ggaagagctc aaggagatca tacagcaagg accagactgg 2040
cttgagctct ctgagatgac cgagcagcca gactatgacc tggagacctt tgtgaacaaa 2100
gcggaatctg ccttgggcca gcaagccaag ctttctcag ccctgcgaga tgtcatcaag 2160
gccttacgcc tggccatgca gctggaagag caggctagca gacaaataag cagcaagaaa 2220
cggccccagt gacgactgca aataaaaaatc tgtttggttt gacaccagc ctcttccttg 2280
gccctcccca gagaactttg ggtacctggt ggggtctaggc agggctctgag ctgggacagg 2340
ttctggtaaa tgccaagtat gggggcatct gggcccaggc cagctgggga gggggtcaga 2400
gtgacatggg acaactcctt tctgttcttc agttgtcgcc ctacagagag gaaggagctc 2460
ttagttaccc ttttgtgttg ccttctcttc catcaagggg aatgttctca gcatagagct 2520
ttctccgcag catcctgcct gcgtggactg gctgctaata gagagctccc tggggttgtc 2580
ctggctctgg ggagagagac ggagccttta gtacagctat ctgctggctc taaaccttct 2640
acgcctttgg gccgagcact gaatgtcttg tactttaaaa aaatgtttct gagacctctt 2700
tctactttac tgtctcccta gagtcttaga ggatccctac tgttttctgt tttatgtgtt 2760
tatacattgt atgtaacaat aaagagaaaa aataaaaaaa aaaaaaaaaa aaaaaaaaaa 2820
aaaaa 2825

```

<210> 216
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006845

<400> 216
 aaatgtttct gagacctctt tctactttac tgtctcccta gagtcttaga ggatccctac 60

<210> 217
 <211> 823
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007019

<400> 217
 aaacgcgggc gggcggggcc gcagtcctgc agttgcagtc gtgttctccg agttcctgtc 60
 tctctgccaa cgccgcccgg atggcttccc aaaaccgcga ccagccgccc actagcgtcg 120
 ccgcccggcg taaaggagct gagccgagcg ggggcgcccgc ccgggggtccg gtgggcaaaa 180
 ggctacagca ggagctgatg accctcatga tgtctggcga taaagggatt tctgccttcc 240
 ctgaatcaga caaccttttc aaatgggtag ggaccatcca tggagcagct ggaacagtat 300
 atgaagacct gaggtataag ctctcgctag agttccccag tggctaccct tacaatgcgc 360
 ccacagtga gttcctcacg ccctgctatc accccaacgt ggacaccag ggtaacatat 420
 gcctggacat cctgaaggaa aagtggctct ccctgtatga tgtcaggacc attctgctct 480
 ccatccagag ctttctagga gaacccaaca ttgatagtcc cttgaacaca catgctgcgc 540
 agctctggaa aaaccccaca gcttttaaga agtacctgca agaaacctac tcaaagcagg 600
 tcaccagcca ggagccctga ccaggtctgc ccagcctgtc cttgtgtcgt ctttttaatt 660
 tttccttaga tggctgttcc tttttgtgat ttctgtatag gactctttat cttgagctgt 720
 ggatattttt ttttgttttt gtcttttaaa ttaagcctcg gttgagccct tgtatattaa 780
 ataaatgcat ttttgtcctt ttttagacaa aaaaaaaaaa aaa 823

<210> 218
 <211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_007019

<400> 218
tggaaaaaacc ccacagcttt taagaagtag ctgcaagaaa cctactcaaa gcaggtcacc 60

<210> 219
<211> 2831
<212> DNA
<213> Homo sapiens

<300>
<308> NM_007183

<400> 219
gaattccgga caggacgtga agatagttgg gtttggaggc ggccgccagg cccaggcccc 60
gtggacctgc cgccatgcag gacggtaact tcctgtctgc ggccctgcag cctgaggccg 120
gcgtgtgctc cctggcgctg ccctctgacc tgcagctgga ccgccggggc gccgaggggc 180
cggaggccga gcggctgcgg gcagcccggg tccaggagca ggtccgcgcc cgcctcttgc 240
agctgggaca gcagccggcg cacaacgggg ccgctgagcc cgagcctgag gccgagactg 300
ccagaggcac atccaggggg cagtaaccaca ccctgcaggc tggcttcagc tctcgctctc 360
agggcctgag tggggacaag acctcggggt tccggcccat cgccaagccg gcctacagcc 420
cagcctcctg gtcctccgcg tccgcctggg atctgagctg cagtcggagg ctgagttcag 480
cccacaatgg gggcagcgcc tttggggccg ctgggtacgg ggggtgccag cccacccctc 540
ccatgccccac caggcccggtg tccctccatg agcgcgggtg ggttgggagc cgggcccact 600
atgacacact ctccctgcgc tcgctgcggc tggggcccgg gggcctggac gaccgctaca 660
gcctgggtgc tgagcagctg gagcccgcgg ccacctccac ctacagggcc tttgcgtacg 720
agcgcaggcc cagctccagc tccagccggg cagggggggt ggactggccc gaggccactg 780
aggtttcccc gagccggacc atccgtgccg ctgcccgtgc gaccctgcag cgattccaga 840
gcagccaccg gagccgcggg gtaggcgggg cagtgcgggg ggccgtcctg gagccagtgg 900
ctcgagcgcc atctgtgcgc agcctcagcc tcagcctggc tgaactgggc cacctgcggg 960
acgtgcatgg gttcaacagc tacggtagcc accgaaccct gcagagactc agcagcgggt 1020
ttgatgacat tgacctgcc tcagcagtca agtacctcat ggcttcagac cccaacctgc 1080
agggtgctggg agcggcctac atccagcaca agtgctacag cgatgcagcc gccaaagaagc 1140
aggcccgag ccttcaggcc gtgcctaggg tgggtgaagc cttcaaccac gccaaaccagg 1200
aagtgcagcg ccatgccaca ggtgccatgc gcaacctcat ctacgacaac gctgacaaca 1260
agctggccct ggtggaggag aacgggatct tcgagctgct ggggacactg cgggagcagg 1320
atgatgagct tcgcaaaaat gtcacaggga tcctgtggaa cctttcatcc agcgaccacc 1380
tgaaggaccg cctggccaga gacacgtcgg agcagctcac ggacctggtg ttgagccccc 1440
tgtcgggggc tgggggtccc cccctcatcc agcagaacgc ctcgaggcg gagatcttct 1500
acaacgccac cggcttcttc aggaacctca gctcagcctc tcaggccact cgccagaaga 1560
tgcgggagtg ccacgggctg gtggacgccc tgggtcacctc tatcaaccac gccctggacg 1620
cgggcaaatg cgaggacaag agcgtggaga acgcgggtgt cgtcctgcgg aacctgtcct 1680
accgcctcta cgacgagatg ccgccgtccg cgtgcagcg gctggagggt cgcggccgca 1740
gggacctggc gggggcgccg ccgggagagg tcgtgggctg cttcacgcgg cagagccggc 1800
ggctgcgcga gctgccctc gccgccgatg cgtcacctt cgcggagggt tccaaggacc 1860
ccaaggccct cgagtggctg tggagcccc agatcgtggg gctgtacaac cggctgctgc 1920
agcgtgcga gctcaaccgg cacacgacgg aggcggccgc cggggcgctg cagaacatca 1980
cggcaggcga ccgcaggtg gcgggggtgc tgagccgcct ggccctggag caggagcgta 2040
ttctgaaccc cctgctagac cgtgtcagga ccgccgacca ccaccagctg cgtcactga 2100
ctggcctcat ccgaaacctg tctcggaacg ctaggaacaa ggacgagatg tccaagaagg 2160
tggtgagcca cctgatcgag aagctgccag gcagcgtggg tgagaagtcg ccccagccg 2220
aggctgtggg caacatcata gctgtgctca acaacctggg ggtggccagc cccatcgctg 2280
cccagacct gctgtatctt gacggactcc gaaagctcat cttcatcaag aagaagcggg 2340
acagccccga cagtgagaag tcctcccggg cagcatccag cctcctggcc aacctgtggc 2400
agtacaacaa gctccaccgt gactttcggg cgaagggtca tcggaaggag gacttcctgg 2460
gcccataggt gaagccttct ggaggagaag gtgacgtggc ccagcgtcca agggacagac 2520
tcagctccag gctgcttggc agcccagcct ggaggagaag gctaatagac gaggggcccc 2580

```

tcgctggggc ccctgtgtgc atctttgagg gtccctggggc accaggagggg gcaggggtctt 2640
atagctgggg acttggcttc cgcaggggcag ggggtggggc aggggtcaag gctgctctgg 2700
tgtatggggg ggtgacccag tcacattggc agagggtggg gttggctgtg gcctggcagt 2760
atcttgggat agccagcact ggaataaag atggccatga acagtcacaa aaaaaaaaaa 2820
aaaaggaatt c 2831

```

<210> 220
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007183

```

<400> 220
ctggcagtat cttgggatag ccagcactgg gaataaagat ggccatgaac agtcacaaaa 60

```

<210> 221
 <211> 2815
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007267

```

<400> 221
aggaagcggg ggaaggtgaa gtaggaccga attcctgtgc cgaagaggcc tgcagtggga 60
gagcaggatg ggggctccgg aggtggcgcc caggctctga gctaccctag gtctgcagac 120
tagcggggcat tggccagaga catggcccag ccactggcct tcatcctoga tgtccctgag 180
accccagggg accaggcca gggccccagc cccatgatg aaagcgaagt gcacgactcc 240
ttccagcagc tcatccagga gcagagccag tgcacggccc aggaggggct ggagctgcag 300
cagagagagc gggaggtgac aggaagtagc cagcagacac tctggcggcc cgagggcacc 360
cagagcacgg ccacactccg catcctggcc agcatgcccc gccgcaccat tggccgcagc 420
cgaggtgcca tcatctccca gtactacaac cgcacggtgc agcttcggtg caggagcagc 480
cgccccctgc tcgggaactt tgtccgctcc gcctggccca gcctccgct gtacgacctg 540
gagctggacc ccacggccct ggaggaggag gagaagcaga gcctcctggt gaaggagtgc 600
cagagcctgg cagtggcaca gcgggaccac atgcttcgcg ggatgccctt aagcctggct 660
gagaaacgca gcctgcgaga gaagagcagg accccgaggg ggaagtggag gggccagccg 720
ggcagcggcg ggggtctgct ctgctgtggc cggctcagat atgcctgcgt gctggccttg 780
cacagcctgg gcctggcgct gctctccgcc ctgcaggccc tgatgccgtg gcgctacgcc 840
ctgaagcgca tcggggggcca gttcggctcc agcgtgctct cctacttcct ctttctcaag 900
accctgctgg ctttcaatgc cctcctgctg ctgctgctgg tggccttcat catggggcct 960
caggctgcct tcccacccgc cctgcgggce cctgcccccg tctgcacagg cctggagctc 1020
ctcacaggcg cgggttgctt caccacacc gtcattgtact acggccacta cagtaacgcc 1080
acgctgaacc agcctgtgtg cagccccctg gatggcagcc agtgcacacc cagggtgggt 1140
ggcctgccct acaacatgcc cctggcctac ctctccactg tgggcgtgag cttctttatc 1200
acctgcatca ccctggtgta cagcatggct cactctttcg gggagagcta ccgggtgggc 1260
agcacctctg gcatccacgc catcaccgtc gcaggacaat attgcacccc ggctgaagga gctgctggcc 1380
aagcgggcct cccgcctcca gcaggacag ccccaggagc gtgtgcggga ggctgcggca ggcggctgtg 1440
gagtggcagc tgcggcacag cccaggagc gtgtctgggg accgcctggt gctgcgccgt gcccgctcac 1500
ctggggcttg tgtggctgct gtgtctgggg cctgctgctg gctgcgccgt gcccgctcac 1560
gtcttctcgg agttcatgat ccagagtcca gaggtgctg gccaggaggc tgtgctgctg 1620
gtcctgcccc tgggtggttg cctcctcaac ctggggggccc cctacctgtg cctgtgctctg 1680
gccgccctgg agccgcatga ctccccggta ctggagggtg acgtggccat ctgcaggaac 1740
ctcatctca agctggccat cctggggaca ctgtgctacc actggctggg ccgcagggtg 1800
ggcgtcctgc agggccagt ctgggaggat tttgtggggc aggagctgta ccggttctctg 1860
gtgatggact tcgtcctcat gttgctggac acgctttttg gggaactggt gtggaggatt 1920
atctccgaga agaagctgaa gaggaggcgg aagccggagt ttgacattgc ccggaatgtc 1980
ctggagctga tttatgggca gactctgacc tggctggggg tgctcttctc gcccctctc 2040
cccgccgtgc agatcatcaa gctgctgctc gtcttctatg tcaagaagac cagccttctg 2100
gccaactgcc aggcgcgcgc ccggccctgg ctggcctcac acatgagcac cgtcttctc

```

```

acgctgctct gcttccccgc cttcctgggc gccgctgtct tcctctgcta cgccgtctgg 2160
caggtgaagc cctcgagcac ctgcggcccc ttccggaccc tggacacccat gtacgaggcc 2220
ggcaggggtgt ggggtgcgcca cctggaggcg gcaggcccca ggggtctcctg gctgccctgg 2280
gtgcaccggt acctgatgga aaacaccttc tttgtcttcc tgggtgtcagc cctgctgctg 2340
gccgtgatct acctcaacat ccaggtgggtg cggggccagc gcaagggtcat ctgcctgctc 2400
aaggagcaga tcagcaatga gggtagggac aaaatcttct taatcaacaa gcttcactcc 2460
atctacgaga ggaaggagag ggaggagagg agcaggggtg ggacaaccga ggaggctgcg 2520
gcacccccctg ccctgctcac agatgaacag gatgcctagg gggacggcga tgggcctcac 2580
gggcccgcgc agcacccctga gaccacactg ttgcctccca gtgacctgc tgggacacca 2640
ggacaaggaa gacagtttcg cctctcgaaa gccgcagctg cgcctaggct ggagctggaa 2700
gggtgggtga atccggcttg ggcaccccca atgaactctg ccctgcctgg gactctattt 2760
attctgatta aaggggtttt gcaaatggga aaaaaaaaaa aaaaaaaaaa aaaaa 2815

```

<210> 222

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007267

<400> 222

```

ggtgaggaca aaatcttctt aatcaacaag cttcactcca tctacgagag gaaggagagg 60

```

<210> 223

<211> 1893

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007274

<400> 223

```

atttaccgcc gcgcgggagag tgaggggcca agtccgccct gctccgccac ttagggcgcc 60
ccagacgctt ccctcggggc tgccaccggg tcgggcggcg ctgccggggc tagcgggcct 120
tcccgcacc ggcgcggccc aaccgccacc gaaccttctg gaagcggcgg ctgcctgggc 180
ccccacgcc ccagaatcgt acgcccgcgc gagctctctg cagccttggc ggcctgggag 240
gcggggctcg ggggtggggc ggcgcggggg cggggtcggc gcggggaggc cgcgttcgat 300
tcgcccccg cgcgcaggcc ccgcctcacc agccccatcg ctccacctct gccctcccc 360
tttatggcgc ggcccgggct cattcattcc gcgcggggc tgccagacac ctgcgccctt 420
ctgcagccgc ccgccgcac cgccgcgcga gccccagca tgtcggggcc agacgtcgag 480
acgccgtccg ccatccagat ctgccggatc atgcggccag atgatgcaa cgtggccggc 540
aatgtccacg gggggaccat cctgaagatg atcgaggagg caggcgccat catcagcacc 600
cggcattgca acagccagaa cggggagcgc tgtgtggccg ccctggctcg tgtcgagcgc 660
accgacttcc tgtctcccat gtgcacgggt gaggtggcgc atgtcagcgc ggagatcacc 720
tacacctcca agcactctgt ggaggtgcag gtcaacgtga tgtccgaaaa catcctcaca 780
ggtgccaaaa agctgaccaa taaggccacc ctgtggatat tgccccctgtc gctgaagaat 840
gtggacaagg tctctgaggt gccctcctgt gtgtattccc ggcaggagca ggaggaggag 900
ggccggaagc ggtatgaagc ccagaagctg gagcgcattg agaccaagtg gaggaacggg 960

```

```

gacatcgctc agccagtcct caaccagag ccgaacactg tcagctacag ccagtccagc 1020
ttgatccacc tgggtggggc ttcagactgc accctgcacg gctttgtgca cggagggtgtg 1080
accatgaagc tcatggatga ggtcgccggg atcgtggctg cagccactg caagaccaac 1140
atcgtcacag cttccgtgga cgcattaat ttcatgaca agatcagaaa aggctgcgtc 1200
atcaccatct cgggacgcat gaccttcacg agcaataagt ccatggagat cgagggtgtg 1260
gtggacgccc accctgttgt ggacagctct cagaagcgt accggggcgc cagtgccttc 1320
ttcacctacg tgtcgctgag ccagggaagg aggtcgctgc ctgtgcccc gctgggtgcc 1380
gagaccgagg acgagaagaa gcgctttgag gaaggcaaag ggcggtacct gcagatgaag 1440
gcgaagcgac agggccacgc ggagcctcag ccctagactc cctcctctg ccactgggtg 1500
ctcagtagtc catggcaacg ggcccagtg ccagtcactt agaagttccc cccttggcca 1560
aaaacccaat tcacattgag agctgggtgt gtctgaagtt ttcgtatcac agtggttaac 1620

```



```

tgtactctct cctgcaaacc tacacaccaa agctttatatt atatcattcc agtatcaatg 1680
ctacacagtg ttgtcccgag cgccgggagg cgttgggcag aaaccctcgg gaatgcttcc 1740
gagcacgctg taggggtatgg gaagaaccca gcaccactaa taaagctgct gcttggctgg 1800
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1860
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 1893

```

```

<210> 224
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_007274

```

```

<400> 224
acctacacac caaagcttta tttatatcat tccagtatca atgctacaca gtgtgtgtccc 60

```

```

<210> 225
<211> 4157
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_007315

```

```

<400> 225
agcggggcgg ggcggccagcg ctgccttttc tcctgcgggg tagtttgcgt ttcctgcgca 60
gagtctgcgg aggggctcgg ctgcaccggg gggatcgcgc ctggcagacc ccagaccgag 120
cagaggcgac ccagcgcgct cgggagaggc tgcaccgccg cgcccccgcc tagcccttcc 180
ggatcctgcg cgcagaaaag tttcatttgc tgtatgccat cctcgagagc tgtctagggt 240
aacgttcgca ctctgtgtat ataacctcga cagtcttggc acctaacgtg ctgtgcgtag 300
ctgctccttt ggttgaatcc ccaggccctt gttggggcac aagggtggcag gatgtctcag 360
tgggtacgaac ttcagcagct tgactcaaaa ttcctggagc aggttcacca gctttatgat 420
gacagttttc ccattggaat cagacagtag ctggcacagt ggtagaaaa gcaagactgg 480
gagcacgctg ccaatgatgt ttcattttgc accatccgtt ttcatgacct cctgtcacag 540
ctggatgac aatatagtcg cttttctttg gagaataact tcttgctaca gcataacata 600
aggaaaagca agcgtaatct tcaggataat tttcagggaag acctaatcca gatgtctatg 660
atcattttaca gctgtctgaa ggaagaaagg aaaattcttg aaaacgcca gagatttaat 720
caggctcagt cggggaatat tcagagcaca gtgatgttag acaaacagaa agagcttgac 780
agtaaagtca gaaatgtgaa ggacaagggt atgtgtatag agcatgaaat caagagcctg 840
gaagatttac aagatgaata tgacttcaaa tgcaaaacct tgcagaacag agaacacgag 900
accaatgggt tggcagaag tgatcagaaa caagaacagc tgttactcaa gaagatgtat 960
ttaatgcttg acaataagag aaaggaagta gttcacaaaa taatagagtt gctgaatgtc 1020
actgaactta ccagaatgc cctgattaat gatgaactag tggagtggaa gccggagacag 1080
cagagcgct gtattggggg gccgcccaat gcttgcttgg atcagctgca gaactgggtc 1140
actatagttg cggagagctt gcagcaagtt cggcagcagc ttaaaaagtt ggaggaattg 1200
gaacagaaat acacctacga acatgacctt atcacaaaaa acaacaagt gttatgggac 1260
cgcaccttca gtcttttcca gcagctcatt cagagctcgt ttgtgggtgga aagacagccc 1320
tgcattgcca cgcacctca gaggccgctg gtcttgaaga caggggtcca gttcactgtg 1380
aagttgagac tgttggtgaa attgcaagag ctgaattata atttgaaagt caaagtctta 1440
tttgataaag atgtgaatga gagaaatata gtaaaaggat ttaggaagtt caacattttg 1500
ggcacgcaca caaaagtgat gaacatggag gagtccacca atggcagctt gccggctgaa 1560
tttcggcacc tgcaattgaa agaacagaaa aatgctggca ccagaacgaa tgagggtcct 1620
ctcatcgtta ctgaagagct tcactccctt agttttgaaa cccaattgtg ccagcctggt 1680
ttggtaattg aacctcgagac gacctctctg ccggttggg tgatctccaa cgtcagccag 1740
ctcccgagcg gttgggcctc catccttttg tacaacatgc tgggtggcga acccaggaat 1800
ctgtccttct tcctgactcc acctgtgca cgtgggctc agctttcaga agtgtctagt 1860
tggcagtttt ctctgtcac caaaagaggt ctcaatgtgg accagctgaa catgttggga 1920
gagaagcttc ttggtcctaa cgccagcccc gatggtctca ttccgtggac gaggttttgt 1980
aaggaaaata taaatgataa aaattttccc ttctggcttt ggattgaaag catcctagaa 2040
ctcattaaaa aacacctgct ccctctctgg aatgatgggt gcatcatggg cttcatcagc 2100

```

```

aaggagcgag agogtgcctt gttgaaggac cagcagccgg ggaccttctt gctgcgggttc 2160
agtgagagct cccgggaagg ggccatcaca ttcacatggg tggagcgggtc ccagaacgga 2220
ggcgaaacctg acttccatgc ggttgaaccc tacacgaaga aagaactttc tgctgttact 2280
ttccctgaca tcattcgcaa ttacaaagtc atggctgctg agaataattc tgagaatccc 2340
ctgaagtatc tgtatccaaa tattgacaaa gaccatgcct ttggaaagta ttactccagg 2400
ccaaaggaag caccagagcc aatggaactt gatggcccta aaggaaactg atatatcaag 2460
actgagttga tttctgtgtc tgaagttcac ccttctagac ttcagaccac agacaacctg 2520
ctccccatgt ctccctgagga gtttgacgag gtgtctcgga tagtggggtc tgtagaattc 2580
gacagtatga tgaacacagt atagagcatg aatttttttc atcttctctg gcgacagttt 2640
tccttctcat ctgtgatttc ctccctgctac tctgttccct cacaatcctg gtttctaggg 2700
aaatgaaaag aaggccagca aattcgctgc aacctgttga tagcaagtga atttttctct 2760
aactcagaaa catcagttac tctgaagggc atcatgcac ttactgaagg taaaattgaa 2820
aggcattctc tgaagagtgg gtttcacaag tgaaaaacat ccagatacac ccaaagtatc 2880
aggacgagaa tgagggtcct ttgggaaagg agaagttaag caacatctag caaatgttat 2940
gcataaagtc agtgcccaac tgttataggt tggttgataa atcagtgggt atttagggaa 3000
ctgcttgacg taggaacggg aaatttctgt gggagaattc ttacatgttt tctttgcttt 3060
aagtgtaacg ggcagttttc cattggttta cctgtgaaat agttcaaagc caagtttata 3120
tacaattata tcagtcctct ttcaaaggta gccatcatgg atctggtagg gggaaaatgt 3180
gtattttatt acatctttca cattggctat ttaaagacaa agacaaattc tgtttcttga 3240
gaagagaata ttagctttac tgtttgttat ggcttaatga cactagctaa tatcaataga 3300
aggatgtaca tttccaaatt cacaagttgt gtttgatata caaagctgaa tacattctgc 3360
tttcatcttg gtcacataca attattttta cagttctccc aaggaggtta ggctattcac 3420
aaccactcat tcaaaagttg aaattaacca tagatgtaga taaactcaga aatttaattc 3480
atgtttctta aatgggctac tttgtccttt ttgttattag ggtggtattt agtctattag 3540
ccacaaaatt gggaaaggag tagaaaaagc agtaactgac aacttgaata atacaccaga 3600
gataatatga gaatcagatc atttcaaaac tcatttccta tgtaactgca ttgagaactg 3660
catatgtttc gctgatatat gtgtttttca catttgcgaa tggttccatt ctctctcctg 3720
tactttttcc agacactttt ttgagtggat gatgtttcgt gaagtatact gtatttttac 3780
ctttttcctt ccttatcact gacacaaaaa gtatattaag agatgggttt gacaagggtc 3840
ttccctttta catactgctg tctatgtggc tgtatcttgt ttttccacta ctgctaccac 3900
aactatatta tcatgcaaat gctgtattct tctttggtgg agataaagat ttcttgagtt 3960
ttgttttaaa attaaagcta aagtatctgt attgcattaa atataatatg cacacagtgc 4020
tttccgtggc actgcataca atctgaggcc tcctctctca gtttttata atagggcgag 4080
aacctaagtt tcagttgatt ttacaattga aatgactaaa aaacaaagaa gacaacatta 4140
aaacaatatt gtttcta 4157

```

<210> 226

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007315

<400> 226

atcagatcat ttcaaaactc atttccatag taactgcatt gagaactgca tatgtttcgc 60

<210> 227

<211> 1696

<212> DNA

<213> Homo sapiens

<300>

<308> NM_009587

<400> 227

```

caaaggactt ctagtggtgt gtgaaaggca gcggtggcca cagaggcggc ggagagatgg 60
ccttcagcgg ttcccaggct ccctacctga gtccagctgt ccccttttct gggactattc 120
aaggaggtct ccaggacgga cttcagatca ctgtcaatgg gaccgttctc agctccagtg 180
gaaccagggt tgctgtgaac tttcagactg gcttcagtgg aaatgacatt gccttccact 240
tcaaccctcg gtttgaagat ggagggtacg tgggtgtgcaa cacgaggcag aacggaagct 300

```

```

ggggggcccgga ggagaggaag acacacatgc ctttccagaa ggggatgccc tttgacctct 360
gcttcctggg gcagagctca gatttcaagg tgatgggtgaa cgggacctc ttcgtgcagt 420
acttccaccg cgtgcccttc caccgtgtgg acaccatctc cgtcaatggc tctgtgcagc 480
tgtcctacat cagcttccag aacccccgca cagtccctgt tcagcctgcc ttctccacgg 540
tgccgttctc ccagcctgtc tgtttccac ccaggcccag ggggcgcaga caaaaacctc 600
ccggcgtgtg gcctgccaac ccggctccca ttaccagac agtcatccac acagtgcaga 660
gcgcccctgg acagatgttc tctactcccg ccattccacc tatgatgtac cccacccccg 720
cctatccgat gcctttcatc accaccattc tgggaggggt gtaccatcc aagtcctacc 780
tctgttcagg cactgtcctg ccagtgctc agaggttcca catcaacctg tgctctggga 840
accacatcgc cttccacctg aacccccgtt ttgatgagaa tgctgtggtc cgcaacaccc 900
agatcgacaa ctctggggg tctgaggagc gaagtctgcc ccgaaaaatg cccttcgtcc 960
gtggccagag cttctcagtg tggatcttgt gtgaagctca ctgcctcaag gtggccgtgg 1020
atggtcagca cctgtttgaa tactaccatc gcctgaggaa cctgccacc atcaacagac 1080
tggaagtggg gggcgacatc cagctgaccc atgtgcagac ataggcgggt tcttggccct 1140
ggggccgggg gctgggggtg ggggcagtc ggtcctctc atcatccca ctccacaggc 1200
ccagcctttc caaccctgcc tgggatctgg gctttaatgc agaggccatg tccttgtctg 1260
gtcctgcttc tggctacagc caccctggaa cggagaaggc agctgacggg gattgccttc 1320
ctcagccgca gcagcacctg gggctccagc tgctggaatc ctaccatccc aggaggcagg 1380
cacagccagg gagagggggg gagtgggcag tgaagatgaa gcccatgct cagtccctc 1440
ccatcccca cgcagctcca cccagtcac aagccaccag ctgtctgctc ctggtgggag 1500
gtggcctcct cagccccctc tctctgacct ttaacctcac tctcaacctg caccgtgcac 1560
caacccttca cccctcctgg aaagcaggcc tgatggcttc ccactggcct ccaccacctg 1620
accagagtgt tctcttcaga ggactggctc ctttccagct gtccttaaaa taaagaaatg 1680
aaaatgcttg ttggca 1696

```

<210> 228

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_009587

<400> 228

```

cagaggactg gctcctttcc cagtgtcctt aaaataaaga aatgaaaatg cttgttggca 60

```

<210> 229

<211> 6552

<212> DNA

<213> Homo sapiens

<300>

<308> NM_012291

<400> 229

```

atgaggagct tcaaaagagt caactttggg actctgctaa gcagccagaa ggaggctgaa 60
gagttgctgc ccgacttgaa ggagttcctg tccaacctc cagctgggtt tccagcagc 120
cgatctgatg ctgagaggag acaagcttgt gatgccatc tgagggttg caaccagcag 180
ctgactgcta agctagcttg ccctaggcat ctggggagcc tgctggagct ggcagagctg 240
gcctgtgatg gctacttagt gtctacccca cagcgtctc cctctacct ggaacgaatt 300
ctctttgtct tactgcggaa tgctgctgca caaggaagcc cagaggccac actccgcctt 360
gctcagcccc tccatgcctg cttggtgcag tgctctcgcg aggtgctcc ccaggactat 420
gaggccgtgg ctcggggcag cttttctctg ctttggagg gggcagaagc cctgttgaa 480
cggcgagctg catttgacg tcggctgaag gccttgagct tctagtaot cttggaggat 540
gaaagtaccc cttgtgaggt tcctcacttt gcttctccaa cagcctgtcg agcggtagct 600
gcccatcagc tatttgatgc cagtggccat ggtctaaatg aagcagatgc tgatttccta 660
gatgacctgc tctccaggca cgtgatcaga gccttggtgg gtgagagagg gagctcttct 720
gggcttcttt ctccccagag ggccctctgc ctottggagc tcaccttggg aactgcccgt 780
cgcttttgct ggagccgcca ccatgacaaa gccatcagcg cagtggagaa ggctcacagt 840
tacctaagga acaccaatct agcccctagc cttcagctat gtcagctggg ggttaagctg 900

```

ctgcaggttg	gggaggaagg	acctcaggca	gtggccaagc	ttctgatcaa	ggcatcagct	960
gtcctgagca	agagtatgga	ggcaccatca	ccccacttc	gggcattgta	tgagagctgc	1020
cagttcttcc	tttcaggcct	ggaacgaggc	accaagaggc	gctatagact	tgatgccatt	1080
ctgagcctct	ttgcttttct	tggagggtac	tgctctcttc	tgcagcagct	gcgggatgat	1140
ggtgtgtatg	ggggctcttc	caagcaacag	cagtcttttc	ttcagatgta	ctttcaggga	1200
cttcacctct	acactgtggg	ggtttatgac	tttgcccaag	gctgtcagat	agttgatttg	1260
gctgacctga	cccaactagt	ggacagttgt	aaatctaccg	ttgtctggat	gctggaggcc	1320
ttagagggcc	tgtcggggca	agagctgacg	gaccacatgg	ggatgaccgc	ttcttacacc	1380
agtaatttgg	cctacagctt	ctatagtcac	aagctctatg	ccgaggcctg	tgccatctct	1440
gagccgctct	gtcagcacct	gggtttgggtg	aagccaggca	cttatcccgga	ggtgcctcct	1500
gagaagttgc	acaggtgctt	ccggctacaa	gtagagagtt	tgaagaaact	gggtaaacag	1560
gccaggggct	gcaagatggg	gattttgtgg	ctggcagccc	tgcaaccctg	tagccctgaa	1620
cacatggctg	agccagtcac	tttctgggtt	cgggtcaaga	tggatgcggc	cagggctgga	1680
gacaaggagc	tacagctaaa	gactctgcga	gacagcctca	gtggctggga	cccggagacc	1740
ctggccctcc	tgtcagggga	ggagctgcag	gectacaagg	cgggtgcgggc	cgacactgga	1800
caggaaacgt	tcaacatcat	ctgtgacctc	ctggagctga	gccccgagga	gacaccagcc	1860
ggggcctggg	cacgagccac	ccacctggta	gaactggctc	aggtgctctg	ctaccacgac	1920
tttacgcagc	agaccaactg	ctctgctctg	gatgctatcc	gggaagccct	gcagcttctg	1980
gactctgtga	ggcctgaggc	ccaggccaga	gatcagcttc	tggacgataa	agcacaggcc	2040
ttgctgtggc	tttacatctg	tactctggaa	gccaaaatac	aggaaggtat	cgagcgggat	2100
cggagagccc	aggccctctg	taacttggag	gaatttgaag	tcaatgacct	gaactatgaa	2160
gataaactcc	aggaagatcg	tttctatata	agtaacattg	ccttcaacct	ggctgcagat	2220
gctgctcagt	ccaaatgcct	ggaccaagcc	ctggccctgt	ggaaggagct	gcttacaagg	2280
gggcaggccc	cagctgtacg	gtgtctccag	cagacagcag	cctcactgca	gatacctagc	2340
gcctcttacc	agctgggtgg	aaagcccatg	caggctctgg	aggtcctcct	gctgctacgg	2400
attgtctctg	agagactgaa	ggaccactcg	aaggcagctg	gctcctcctg	ccacatcacc	2460
cagctcctcc	tgaccctcgg	ctgtcccagc	tatgcccagt	tacacctgga	agaggcagca	2520
tcgagcctga	agcatctcga	tcagactact	gacacatacc	tgctcctttc	cctgacctgt	2580
gatctgcttc	gaagtcaact	ctactggact	caccagaagg	tgaccaaggg	tgctctctct	2640
ctgctgtctg	tgcttcggga	tcttgccttc	cagaagtcct	ccaaggcttg	gtacttgctg	2700
cgtgtccagg	tcctgcagct	ggtggcagct	taccttagcc	tcccgtaaaa	caacctctca	2760
cactcctgt	gggagcagct	ctgtgcccaa	ggctggcaga	caacctgagat	agctctcata	2820
gactcccata	agctcctccg	aagcatcatc	ctcctgctga	tgggcagtga	cattctctca	2880
actcagaaag	cagctgtgga	gacatcgttt	ttggactatg	gtgaaaatct	ggtacaaaaa	2940
tggcaggttc	tttcagaggt	gctgagctgc	tcagagaagc	tggcttgcca	cctgggccgc	3000
ctgggtagt	tgagtgaagc	caaggccttt	tgcttggagg	ccctaaaact	tacaacaaag	3060
ctgcagatac	cacgccagtg	tgccctgttc	ctggtgctga	agggcgagct	ggagctggcc	3120
cgcaatgaca	ttgatctctg	tcagtcggac	ctgcagcagg	ttctgttctt	gcttgagctc	3180
tgcacagagt	ttggtggggg	gactcagcac	ctggactctg	tgaagaaggt	ccacctgcag	3240
aaggggaagc	agcaggccca	ggtccctgt	ctccacacgc	tcccagagga	ggagctcttc	3300
ctaagaggcc	ctgctctaga	gctgggtggc	actgtggcca	aggagcctgg	ccccatgca	3360
ccttctacaa	actcctcccc	agtcttgaaa	accaagcccc	agcccatacc	caacttctct	3420
tcccattcac	ccacctgtga	ctgctcgctc	tgcgccagcc	ctgtcctcac	agcagctctg	3480
ctgcgcctgg	tattgggtcac	ggcaggggtg	aggctggcca	tgggccacca	agcccagggt	3540
ctggatctgc	tgcaggctgt	gctgaagggc	tgtcctgaag	ccgctgagcg	cctcacccaa	3600
gctctccaag	cttccctgaa	tcataaaaac	ccccctcct	tgggttccaag	cctcttggat	3660
gagatcttgg	ctcaagcata	cacactgttg	gcactggagg	gcttgaacca	gccatcaaac	3720
gagagcctgc	agaaggttct	acagtcaggg	ctgaagtttg	tagcagcacg	gataccccac	3780
ctagagccct	ggcgagccag	cctgctcttg	atttggggcc	tcacaaaact	aggtggcctc	3840
agctgctgta	ctacccaact	ttttgcaagc	tcttggggct	ggcagccacc	attaataaaa	3900
agtgtccctg	gctcagagcc	ctctaagact	cagggccaaa	aacgttcttg	acgagggcgc	3960
caaaagttag	cctctgctcc	cctgcgcctc	aataatacct	ctcagaaagg	tctggaaggt	4020
agagctatgc	cctgcacacc	taaaccccca	gaccggatca	ggcaagctgg	ccctcatgtc	4080
cccttcacgg	tgtttgagga	agtctgccct	acagagagca	agcctgaagt	accccagccc	4140
cccaggggtac	aacagagagt	ccagacgcgc	ctcaaggtga	acttcagtga	tgacagtgc	4200
ttggaagacc	ctgtctcagc	tgaggcctgg	ctggcagagg	agcctaagag	acggggcact	4260
gcttcccggg	gcccggggcg	agcaagggaag	ggcctgagcc	taaagacgga	tgccgtgggt	4320
gccccaggta	gtgcccctgg	gaaccctggc	ctgaatggca	ggagccggag	ggccaagaag	4380
gtggcatcaa	gacattgtga	ggagcggcgt	ccccagaggg	ccagtgacca	ggccaggcct	4440

```

ggccctgaga tcatgaggac catccctgag gaagaactga ctgacaactg gagaaaaatg 4500
agcttttgaga tctcagggg ctctgacggg gaagactcag cctcaggtgg gaagactcca 4560
gctccggggcc ctgaggcagc ttctggagaa tgggagctgc tgaggctgga ttccagcaag 4620
aagaagctgc ccagcccatg cccagacaag gagagtgaca aggacctggg tcctcggctc 4680
cagctccccct cagccccctg agccactggt ctttctaccc tggactccat ctgtgactcc 4740
ctgagtgttg ctttccgggg cattagtcac tgtcctccta gtgggctcta tgcccacctc 4800
tgccgcttcc tggccttggt cctggggccac cgggacccct atgccactgc tttccttgtc 4860
accgagctcg tctccatcac ctgtcgccac cagctgctca cccacctcca cagacagctc 4920
agcaaggccc agaagcaccg aggatcactt gaaatagcag accagctgca ggggctgagc 4980
cttcaggaga tgccctggaga tgtcccccct gcccgcctcc agcgcctctt ttccttcagg 5040
gcttttgaat ctggccactt ccccagcct gaaaaggaga gtttccagga gcgcctggct 5100
ctgatcccca gtggggtgac tgtgtgtgtg ttggccctgg ccacctcca gcccggaacc 5160
gtgggcaaca cctcctgct gaccggctg gaaaaggaca gtccccagc cagtgtgcag 5220
attccactg gccagaacaa gcttcatctg cgttcagtcc tgaatgagtt tgatgccatc 5280
cagaaggcac accacaggat ggaggttctc atcgcttccc tagagaagtc tgtgctgggc 5340
ctggcactgg accacaggat ggaggttctc atcgcttccc tagagaagtc tgtgctgggc 5400
tgctggaagg ggctgctgct gccgtccagt gaggagcccg gccctgcca ggaggcctcc 5460
cgctacagg agctgctaca ggactgtggc tggaaatatc ctgaccgcac tctgctgaaa 5520
atcatgctca gtggtgcccg tgccctcacc cctcaggaca ttcaggccct ggcctacggg 5580
ctgtgcccga cccagccaga gcgagcccg gagctcctga atgaggcagt aggacgtcta 5640
cagggcctga cagtaccaag caatagccac cttgtcttgg tcctagacaa ggacttgca 5700
aagctgccg gccgaaagcat gccagcctc caagcactgc ctgtcaccg gctgccctcc 5760
ttccgcttcc tactcagcta ctccatcatc aaagagtatg gggcctcgcc agtgctgagt 5820
caaggggtgg atccacgaag taccttctat gtctgaacc ctcaacaata cctgtcaagc 5880
acagaggagc aatttcgagc caatttcagc agtgaagctg gctggagagg agtgggtggg 5940
gaggtgccaa gacctgaaca ggtgcaggaa gccctgacaa agcatgattt gtatatctat 6000
gcagggcatg gggctggtgc ccgcttccct gatgggcagg ctgtcctgcg gctgagctgt 6060
cgggcagtg ccctgctgtt tggctgtagc agtgcgccc tggctgtgca tggaaacctg 6120
gagggggctg gcatcgctgc caagtacatc atggctggtt gcccttggtt tctgggtaat 6180
ctctgggatg tgactgaccg cgacattgac cgctacacgg aagctctgct gcaaggctgg 6240
cttgagcag gccaggggc ccccttctc tactatgtaa accaggcccg ccaagctccc 6300
cgactcaagt atcttattgg ggctgcacct atagcctatg gcttgctgt ctctctgagg 6360
taaccccatg gagctgtctt attgatgcta gaagcctcat aactgttcta cctccaagg 6420
tagatttaac ccttaggata actcttttaa agtgattttc ccagtggtt tatatgaaac 6480
atttcctttt gatttaacct cagtataata aagatacatc atttaaacc tgaaaaaaa 6540
aaaaaaaaaa aa 6552

```

<210> 230

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_012291

<400> 230

```

agcctcataa ctgttctacc tccaagggtta gatttaatcc ttaggataac tcttttaaaag 60

```

<210> 231

<211> 6317

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013261

<400> 231

```

tagtaagaca ggtgccttca gttcactctc agtaaggggc tgggtgacct catgagtgtg 60
tgctctgtgt cactgtggat tggagttgaa aaagcttgac tggcgctcatt caggagctgg 120

```

atggcggtggg	acatgtgcaa	ccaggactct	gagtctgtat	ggagtgcacat	cgagtgtgct	180
gctctggttg	gtgaagacca	gcctctttgc	ccagatcttc	ctgaacttga	tctttctgaa	240
ctagatgtga	acgacttgga	tacagacagc	tttctgggtg	gactcaagtg	gtgcagtgc	300
caatcagaaa	taatattcaa	tcagtacaac	aatgagcctt	caaacatatt	tgagaagata	360
gatgaagaga	atgaggcaaa	cttgctagca	gtcctcacag	agacactaga	cagtctccct	420
gtggatgaag	acggattgcc	ctcatttgat	gcgctgacag	atggagacgt	gaccactgac	480
aatgaggcta	gtccttcctc	catgcctgac	ggcacccttc	cacccagga	ggcagaagag	540
ccgtctctac	ttaagaagct	cttactggca	ccagccaaca	ctcagctaag	ttataatgaa	600
tgcagtggtc	tcagtaccca	gaacctgca	aatcacaatc	acaggatcag	aacaaacctt	660
gcaattgtta	agactgagaa	ttcatggagc	aataaagcga	agagtatttg	tcaacagcaa	720
aagccacaaa	gacgtccctg	ctcggagctt	ctcaaataatc	tgaccacaaa	cgatgaccct	780
cctcacacca	aaccacacaga	gaacagaaac	agcagcagag	acaaatgcac	ctccaaaaag	840
aagtcccaca	cacagtgcga	gtcacaacac	ttacaagcca	aaccaacaac	tttatctctt	900
cctctgaccc	cagagtccac	aaatgacccc	aaggggttccc	catttgagaa	caagactatt	960
gaacgcacct	taagtgtgga	actctctgga	actgcaggcc	taactccacc	caccactcct	1020
cctcataaag	ccaaccaaga	taaccttttt	agggcttctc	caaagctgaa	gtcctcttgc	1080
aagactgttg	tgccaccacc	atcaaagaa	cccaggtaca	gtgagcttct	tggtacacaa	1140
ggcaataact	ccaccaagaa	agggccggag	caatccgagt	tgtatgcaca	actcagcaag	1200
tcctcagtc	tcactgggtg	acacgaggaa	aggaagacca	agcgggccag	tctgcggctg	1260
tttggtgacc	atgactattg	ccagtcgaat	aattccaaaa	cagaaatact	cattaatata	1320
tcacaggagc	tccaagactc	tagacaacta	gaaaataaa	atgtctcttc	tgattggcag	1380
gggcagattt	gttcttccac	agattcagac	cagtgctacc	tgagagagac	tttggaggca	1440
agcaagcagg	tctctccttg	cagcacaaga	aaacagctcc	aagaccagga	aatccgagcc	1500
gagctgaaca	agcacttcgg	tcattcccagt	caagctgttt	ttgacgacga	agcagacaag	1560
accggtgaac	tgaggggacag	tgatttcagt	aatgaacaat	tctccaaact	acctatgttt	1620

ataaattcag	gactagccat	ggatggcctg	tttgatgaca	gcgaagatga	aagtgataaa	1680
ctgagctacc	cttgggatgg	cacgcaatcc	tattcattgt	tcaatgtgtc	tccttcttgt	1740
tcttctttta	actctccatg	tagagattct	gtgtcaccac	ccaaatcctt	atcttctcaa	1800
agaccccaaa	ggatgcgctc	tcgttcaagg	tccttttctc	gacacagggtc	gtgttcccg	1860
tcaccatatt	ccaggtcaag	atcaaggctc	ccaggcagta	gatcctcttc	aagatcctgc	1920
tattactatg	agtcaagcca	ctacagacac	cgcacgcacc	gaaattctcc	cttgatgtg	1980
agatcacgtt	caagatcgcc	ctacagccgt	cggcccaggt	atgacagcta	cgaggaatat	2040
cagcacgaga	ggctgaagag	ggaagaatat	cgcagagagt	atgagaagcg	agagtctgag	2100
agggccaagc	aaagggagag	gcagaggcag	aaaggcaattg	aagagcgccg	tgtgatttat	2160
gtcggtaaaa	tcagacctga	cacaacacgg	acagaactga	gggaccgttt	tgaagttttt	2220
ggtgaaattg	aggagtgcac	agtaaatctg	cgggatgatg	gagacagcta	tggtttcatt	2280
acctaccgtt	atacctgtga	tgcttttgct	gctcttgaaa	atggatacac	tttgccgagg	2340
tcaaacgaaa	ctgactttga	gctgtacttt	tgtggacgca	agcaattttt	caagtctaac	2400
tatgcagacc	tagattcaaa	ctcagatgac	tttgaccctg	cttccaccaa	gagcaagtat	2460

gactctctgg	atcttgatag	tttactgaaa	gaagctcaga	gaagcttgog	caggtaacat	2520
gttccttagc	tgaggatgac	agagggatgg	cgaatacctc	atgggacagc	gcgtccttcc	2580
ctaaagacta	ttgcaagtca	tacttaggaa	tttctcctac	tttactctct	ctgtacaaaa	2640
acaaaacaaa	acaacaacaa	tacaacaaga	acaacaacaa	caataacaac	aatggtttac	2700
atgaacacag	ctgctgaaga	ggcaagagag	agaatgatat	ccagtaagca	catgtttatt	2760
catgggtgtc	agctttgctt	ttcctggagt	ctcttggtga	tggagtgtgc	gtgtgtgcat	2820
gtatgtgtgt	gtgtatgtat	gtgtgtgggtg	tgtgtgtgtg	gttttagggga	agtatgtgtg	2880
ggtacatgtg	aggactgggg	gcacctgacc	agaatgcgca	agggcacaacc	atttcaaatg	2940
gcagcagttc	catgaagaca	cgcttaaaac	ctagaacttc	aaaatgttcg	tattctattc	3000
aaaaggaaat	atatatatat	atatatatat	atatatatat	atatataaat	taaaaaggaa	3060
agaaaactaa	caaccaacca	accaaccaac	caaccacaaa	ccaccctaaa	atgacagccg	3120
ctgatgtctg	ggcatcagcc	tttgtactct	gttttttttaa	gaaagtgcag	aatcaacttg	3180
aagcaagctt	tctctcataa	cgtaatgatt	atatgacaat	cctgaagaaa	ccacaggttc	3240
catagaacta	atatcctgtc	tctctctctc	tctctctctc	tctctttttt	ttttcttttt	3300
ccttttgcca	tggaatctgg	gtgggagagg	atactgcggg	caccagaatg	ctaaagtttc	3360
ctaacatttt	gaagtttctg	tagttcatcc	ttaatcctga	cacccatgta	aatgtccaaa	3420
atgttgatct	tccactgcaa	atttcaaaa	ccttgtcaat	ggtcaagcgt	gcagcttggt	3480
cagcggttct	ttctgaggag	cggacaccgg	gttacattac	taatgagagt	tgggtagaac	3540
tctctgagat	gtgttcagat	agtgtaatg	ctacattctc	tgatgtagtt	aagtatttac	3600

```

agatgttaaa tggagtat tttttttatg tatatactat acaacaatgt tcttttttgt 3660
tacagctatg cactgtaaat gcagccttct tttcaaaact gctaaatttt tcttaataca 3720
gaatattcaa atgtaattat gaggtgaaac aattattgta cactaacata tttagaagct 3780
gaacttactg cttatatata tttgattgta aaaacaaaaa gacagtgtgt gtgtctgttg 3840
agtgcaacaa gagcaaaatg atgctttccg cacatccatc ccttaggtga gcttcaatct 3900
aagcatcttg tcaagaaata tcctagtccc ctaaagggtat taaccacttc tgcgatattt 3960
ttccacattt tcttgtcgct tgtttttctt tgaagtttta tacactggat ttgttagggg 4020
aatgaaattt tctcatctaa aattttttcta gaagatatca tgattttatg taaagtctct 4080
caatgggtaa ccattaagaa atgtttttat tttctctatc aacagtagtt ttgaaactag 4140
aagtcaaaaa tctttttaaa atgctgtttt gttttaattt ttgtgatttt aatttgatac 4200
aaaatgctga ggtaataatt atagtatgat ttttacaata attaatgtgt gtctgaagac 4260
tatctttgaa gccagtattt ctttcccttg gcagagtatg acgatgggtat ttatctgtat 4320
tttttacagt tatgcatcct gtataaatac tgatatttca ttcctttgtt tactaaagag 4380
acataatttat cagttgcaga tagcctattt attataaatt atgagatgat gaaaataata 4440
aagccagtgg aaattttcta cctaggatgc atgacaattg tcaggttgga gtgtaagtgc 4500
ttcatttggg aaatttcagct tttgcagaag cagtgtttct acttgcacta gcatggcctc 4560
tgacgtgacc atgggtgtgt tcttgatgac attgcttctg cttaaattta taaaaacttc 4620
agaaaaacct ccattttgat catcaggatt tcatctgagt gtggagtcctc tggaaatggaa 4680
ttcagtaaca tttggagtgt gtattcaagt ttctaaattg agattcgatt actgtttggc 4740
tgacatgact tttctggaag acatgatata cctactactc aattgttctt ttcctttctc 4800
tcgccaacaa cgatcttgta agatggattt cacccccagg ccaatgcagc taattttgat 4860
agctgcattc atttatcacc agcatattgt gttctgagtg aatccactgt ttgtcctgtc 4920
ggatgcttgc ttgatttttt ggcttcttat ttctaagtag atagaaagca ataaaaatac 4980
tatgaaatga aagaacttgt tcacagggtc tgcgttaca cagtaacaca tctttaatcc 5040
gcctaattct tgttgttctg taggttaaat gcagggtatt taactgtgtg aacgccaac 5100
taaagtttac agtctttctt tctgaatttt gagtatcttc tgtttagtaa taataataaa 5160
aagactatta agagcaataa attattttta agaaatcgag atttagtaaa tccattatg 5220
tgttcaagga ccacatgtgt tctctatttt gccttttaaat ttttgtgaac caatttttaa 5280
tacattctcc tttttgccct ggattgttga catgagtggg atacttgggt tcttttctta 5340
cttatcaaaa gacagcacta cagatatcat attgaggatt aatttatccc cctaccccc 5400
agcctgacaa atattgttac catgaagata gttttcctca atggacttca aattgcatct 5460
agaattagtg gagcttttgt atcttctgca gacactgtgg gtagcccatc aaaatgtaag 5520
ctgtgctcct ctcatttttt tttttatttt tttgggagag aatatattca atgaacacgt 5580
gcaccccatc atcactggag gcaaatttca gcatagatct gtaggatttt tagaagaccg 5640
tgggccattg ccttcatgcc gtggtaagta ccacatctac aattttggta accgaactgg 5700
tgcttttagta atgtggattt ttttcttttt taaaagagat gtagcagaat aattcttcca 5760
gtgcaacaaa atcaattttt tgctaaacga ctccgagaa aacagttggg ctgtcaacat 5820
tcaaagcagc agagagggaa ctttgcacta ttgggggtat atgtttgggt cagttgataa 5880
aaggaaacct tttcatgcct ttagatgtga gcttccagta ggtaatgatt atgtgtcctt 5940
tcttgatggc tgtaatgaga acttcaatca ctgtagtcta agacctgatc tatagatgac 6000
ctagaatagc catgtactat aatgtgatga ttctaaattt gtacctatgt gacagacatt 6060
ttcaataatg tgaactgctg atttgatgga gctactttta gattttagg tgaaagtgt 6120
atactgttgg ttgaactatg ctgaagaggg aaagtgagcg attagttag cccttgccgg 6180
gccttttttc cacctgccaa ttctacatgt attgttgtgg ttttattcat tgtatgaaaa 6240
ttcctgtgat tttttttaaa tgtgcagtac acatcagcct cactgagcta ataaagggaa 6300
acgaatgttt caaatct 6317

```

<210> 232

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013261

<400> 232

ctgtagtcta agacctgatc tatagatacc tagaatagcc atgtactata atgtgatgat 60

<210> 233

<211> 3237

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013277

<400> 233

gcgaagtga	gggtggccca	ggtggggcca	ggctgactga	atgtatctcc	tagctatgga	60
ctaaataata	catgggggga	aataaacaag	tattcatgag	ggtgaaaatg	tgaccagca	120
ggaaaattac	aactattttc	aattgacgtt	gaataggatg	agtcatggaa	tttaagtgat	180
ttactgaaga	ttatactact	ggtagataga	agagctaaag	aaagatggat	actatgatgc	240
tgaatgtgcg	gaatctgttt	gagcagcttg	tgcgccgggt	ggagattctc	agtgaaggaa	300
atgaagtcca	atztatccag	ttggcgaagg	actttgagga	tttccgtaaa	aagtggcaga	360
ggactgacca	tgagctgggg	aaatacaagg	atcttttgat	gaaagcagag	actgagcgaa	420
gtgctctgga	tgtaagctg	aagcatgcac	gtaatcaggt	ggatgtagag	atcaaacgga	480
gacagagagc	tgaggctgac	tgcgaaaagc	tggaacgaca	gattcagctg	attcgagaga	540
tgtctcatgtg	tgacacatct	ggcagcattc	aactaagcga	ggagcaaaaa	tcagctctgg	600
ctttttctcaa	cagaggccaa	ccatccagca	gcaatgctgg	gaacaaaaga	ctatcaacca	660
ttgatgaatc	tggttccatt	ttatcagata	tcagctttga	caagactgat	gaatcactgg	720
attgggactc	ttctttgggtg	aagactttca	aactgaagaa	gagagaaaag	agggcgtcta	780
ctagccgaca	gtttgttgat	ggtccccctg	gacctgtaaa	gaaaactcgt	tccattggct	840
ctgcagtaga	ccaggggaat	gaatccatag	ttgcaaaaac	tacagtgact	gttcccaatg	900
atggcggggc	catcgaagct	gtgtccacta	ttgagactgt	gccatattgg	accaggagcc	960
gaaggaaaac	aggtacttta	caaccttggg	acagtactc	cacctgaac	agcaggcagc	1020
tggagccaag	aactgagaca	gacagtgtgg	gcacgccaca	gagtaatgga	gggatgcgcc	1080
tgcattgactt	tgtttctaag	acggttatta	aacctgaatc	ctgtgttcca	tgtggaaagc	1140
ggataaaaatt	tggcaaatga	tctctgaagt	gtcgagactg	tcgtgtgggtc	tctcatccag	1200
aatgtcggga	ccgctgtccc	cttccctgca	ttcctaccct	gataggaaca	cctgtcaaga	1260
ttggagaggg	aatgctggca	gactttgtgt	cccagacttc	tccaatgato	ccctccattg	1320
ttgtgcattg	tgtaaatgag	attgagcaaa	gaggtctgac	tgagacaggc	ctgtatagga	1380
tctctggctg	tgaccgcaca	gtaaaagagc	tgaaagagaa	attcctcaga	gtgaaaactg	1440
tacccctcct	cagcaaagtg	gatgatatcc	atgctatctg	tagccttcta	aaagactttc	1500
ttcgaaacct	caaagaacct	cttctgacct	ttcgccttaa	cagagccttt	atggaagcag	1560
cagaaatcac	agatgaagac	aacagcatag	ctgccatgta	ccaagctggt	ggtgaactgc	1620
cccaggccaa	cagggaacaca	ttagctttcc	tcattgattca	cttgacagaga	gtggctcaga	1680
gtccacatac	taaaatggat	gttgccaatc	tggctaaagt	ctttggccct	acaatagtgg	1740
cccatgctgt	gcccaatcca	gacccagtga	caatgtttaca	ggacatcaag	cgtcaaccca	1800
aggtgggtga	gcgcctgctt	tccttgccctc	tggagtattg	gagtcagttc	atgatggtgg	1860
agcaagagaa	cattgacccc	ctacatgtca	ttgaaaactc	aaatgccttt	tcaacaccac	1920
agacaccaga	tattaaagtg	agtttactgg	gacctgtgac	cactcctgaa	catcagcttc	1980
tcaagactcc	ttcatctagt	tcctgtgtcac	agagagctccg	ttccaccctc	accaagaaca	2040
ctcctagatt	tgggagcaaa	agcaagtctg	ccagtaacct	aggacgacaa	ggcaactttt	2100
tgtcttctoc	aatgctcaag	tgaagtcaca	tctgcctgtt	acttcccagc	attgactgac	2160
tataagaaag	gacacatctg	tactctgtctc	tgcagcctcc	tgtactcatt	actactttta	2220
gcattctcca	ggcttttact	caagtttaaat	tgtgcatgag	ggttttatta	aaactatata	2280
tatctccctc	tccttctcct	caagtcacat	aatatcagca	ctttgtgctg	gtcattgttg	2340
ggagcttttta	gatgagacat	ctttccaggg	gtagaagggg	tagtatggaa	ttggttgtga	2400
ttcttttttg	ggaagggggg	tattgttctc	ttggcttaaa	gccaaatgct	gctcatagaa	2460
tgatctttct	ctagtttcat	ttagaactga	tttccgtgag	acaatgacag	aaaccctacc	2520
tatctgataa	gattagcttg	tctcaggggtg	ggaagtggga	gggcagggca	aagaaaggat	2580
tagaccagag	gatttaggat	gcctccttct	aagaaccaga	agttctcatt	ccccattatg	2640
aactgagcta	taatattggag	ctttcataaa	aatgggatgc	attgaggaca	gaactagtga	2700
tgggagtatg	cgtagctttg	atttggatga	ttaggtcttt	aatagtgttg	agtggcacaa	2760
ccttgtaaat	gtgaaagtac	aactcgtatt	tatctctgat	gtgccgctgg	ctgaactttg	2820
ggttcatttg	gggtcaaaagc	cagtttttct	tttaaaattg	aattcattct	gatgcttggc	2880
ccccatcccc	ccaaccttgt	ccagtggagc	ccaacttcta	aagggtcaata	tatcatcctt	2940
tggcatccca	actaacaata	aagagtaggc	tataagggga	gattgtcaat	attttgtggg	3000
aagaaaagct	acagtcattt	tttcttttga	ctttggatgc	tgaaattttt	cccatggaa	3060
atagccacat	ctagatagat	gtgagctttt	tcttctgtta	aaattattct	taatgtctgt	3120
aaaaacgatt	ttcttctgta	gaatgtttga	cttcgtattg	acccttatct	gtaaaacacc	3180

tattttgggat aatattttgga aaaaaagtaa atagctttttt caaaatgaaa aaaaaaa 3237

<210> 234

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013277

<400> 234

ctcattcccc attatgaact gagctataat atggagctttt cataaaaaatg ggatgcattg 60

<210> 235

<211> 1122

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013409

<400> 235

```
gctcctcgcc ccgcgcctgc ccccaggatg gtccgcgcga ggcaccagcc ggggtgggctt 60
tgccctctgc tgctgctgct ctgccagttc atggaggacc gcagtgccca ggctgggaac 120
tgctggctcc gtcaagcgaa gaacggccgc tgccaggctc tgtacaagac cgaactgagc 180
aaggaggagt gctgcagcac cggccggctg agcacctcgt ggaccgagga ggacgtgaat 240
gacaacacac tcttcaagtg gatgattttc aacgggggcg cccccaactg catccctgt 300
aaagaaaacgt gtgagaacgt ggactgtgga cctgggaaaa aatgccgaat gaacaagaag 360
aacaaccccc gctgcgtctg cgcgccggat tgttccaaca tcacctgaa ggggtccagtc 420
tgccggctgg atgggaaaac ctaccgcaat gaatgtgcac tcctaaaggc aagatgtaaa 480
gagcagccag aactggaagt ccagtaocaa ggcagatgta aaaagacttg tcgggatgtt 540
ttctgtccag gcagctccac atgtgtggtg gaccagacca ataatgccta ctgtgtgacc 600
tgtaatcgga tttgcccaga gcctgcttcc tctgagcaat atctctgtgg gaatgatgga 660
gtcacctact ccagtgcctg ccacctgaga aagctacct gectgctggg cagatctatt 720
ggattagcct atgagggaaa gtgtatcaaa gcaaagtccg gtgaagatat ccagtgcact 780
gggtgggaaaa aatgtttatg ggatttcaag gttgggagag gccggtgttc cctctgtgat 840
gagctgtgcc ctgacagtaa gtcggatgag cctgtctgtg ccagtgacaa tgccacttat 900
gccagcgagt gtgccatgaa ggaagctgcc tgctcctcag gtgtgctact ggaagtaaag 960
cactccggat cttgcaactc catttcggaa gacaccgagg aagaggagga agatgaagac 1020
caggactaca gctttcctat atcttotatt ctagagtggg aaactctcta taagtgttca 1080
gtgttcacat agcctttgtg caaaaaaaaa aaaaaaaaaa aa 1122
```

<210> 236

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013409

<400> 236

gaagatgaag accaggacta cagcttttct atatcttcta ttctagagtg gtaaactctc 60

<210> 237

<211> 11389

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014246

<400> 237

atggcgccgc	cgccgcgcgc	cgtgctgccc	gtgctgctgc	tccctggccgc	cgccgcgcgc	60
ctgccggcga	tggggctgcg	agcggccgcg	tgggagccgc	gcgtacccgc	cgggaccgcg	120
gccttcgccc	tccggcccgc	ctgtacctac	gcggtgggcg	ccgcttgca	gccccgggcg	180
ccgcgggagc	tgctggacgt	gggcgcgcgt	gggcggctgg	caggacgtcg	gcgcgtctcg	240
ggcgccgggg	gcccgcgtgc	gctgcaagtc	cgcttggtgg	ccgcagtgcc	ccgcagggcg	300
ctgagccgcc	gcctgcccgc	gcgcacgcac	cttcccggct	gcggagcccg	tgccccggctc	360
tgccgaaccg	gtgcccggct	ctgcggggcg	ctctgcttcc	ccgtccccgc	cggtcgccgc	420
gcccgcgagc	attcgggcgt	cgcagctccg	accaccttac	ccgcctgccc	ctgcccgcgc	480
cgccccaggc	cccgtgttcc	cgcccgctcc	atctgtcctg	cgccggggcg	ctcggtccgc	540
ctgcgtctgc	tgtgcgccct	gcggcgccgc	gctggcgccg	tccgggtggg	actggcgctg	600
gaggccgcca	ccgcggggac	gcccctccgc	tcccatccc	catcgccgc	cctgccgcgc	660
aacttgcccc	aagcccgggc	ggggccggcg	cgacgggccc	ggcggggcac	gagcggcaga	720
gggagccctga	agtttccgat	gcccactac	caggtggcgt	tgtttgagaa	cgaaccggcg	780
ggcaccctca	tccctccagct	gcacgcgcac	tacaccatcg	agggcgagga	ggagcgcgtg	840
agctattaca	tggaggggct	gttcgacgag	cgctcccggg	gctacttccg	aatcgactct	900
gccacggggc	ccgtgagcac	ggacagcgta	ctggaccgcg	agaccaagga	gacgcacgtc	960
ctcaggggtga	aagccgtgga	ctacagtacg	ccgcgcgcgt	cgccaccac	ctacatcact	1020
gtcttggtca	aagacaccaa	cgaccacagc	ccggtcttcg	agcagtcgga	gtaccgcgag	1080
cgcggtgcggg	agaacctgga	ggtgggctac	gaggtgctga	ccatccgcgc	cagcgaccgc	1140
gactcgccca	tcaacgccaa	cttgcgttac	cgcggtgttg	ggggcgcgctg	ggacgtcttc	1200
cagctcaacc	agagctctgg	cgtgggtgagc	acacggggcg	tgctggaccg	ggaggaggcg	1260
gccgagtacc	agctcctggg	ggaggccaac	gaccaggggc	gcaatccggg	cccgtcagct	1320
gccacggcca	ccgtgtacat	cgaggtggag	gacgagaacg	acaactaccc	ccagttcagc	1380
gagcagaact	acgtgggtca	ggtgcccag	gacgtggggc	tcaacacggc	tgtgctgcga	1440
gtgcaggcca	cggaccggga	ccagggccag	aacgcggcca	ttcactacag	catcctcagc	1500
gggaacgtgg	ccggccagtt	ctacctgcac	tgcgtgagcg	ggatcctgga	tgtgatcaac	1560
cccttggtatt	tcagggatgt	ccagaaatac	tgcgtgagca	ttaaggccca	ggatgggggc	1620
cggccccgcg	tccaggggtg	ttcaggggtg	gtgtctgtgc	aggtgctgga	tgtcaacgac	1680
aacgagccta	tctttgtgag	cagcccttcc	caggccacgg	tgctggagaa	tgtgcccctg	1740
ggctaccccg	tggtgcacat	tcaggcggtg	gacgcggact	ctggagagaa	cgcccggtcg	1800
cactatcgcc	tggtggacac	ggcctccacc	tttctggggg	gcggcagcgc	tgggccctaag	1860
aatcctgccc	ccacctctga	cttccccttc	cagatccaca	acagctccgg	ttggatcaca	1920
gtgtgtgccg	agctggaccg	cgaggagggt	gagcactaca	gcttcggggg	ggaggcggtg	1980
gaccacggct	cgcccccat	gagctcctcc	accagcgtgt	ccatcacggg	gctggacgtg	2040
aatgacaacg	acccggtgtt	cacgcagccc	acctacgagc	ttcgtctgaa	tgaggatgcg	2100
gccgtgggga	gcagcgtgct	gacctgcag	gcccgcgacc	gtgacgcca	cagtgtgatt	2160
acctaccagc	tcacaggcgg	caacacccgg	aaccgctttg	cactcagcag	ccagagaggg	2220
ggcggcctca	tcacctggc	gctacctctg	gactacaagc	aggagcagca	gtacgtgctg	2280
gcggtgacag	catccgacgg	cacacggctc	cacactgcgc	atgtccta	caacgtcact	2340
gatgccaaca	cccacaggcc	tgtctttcag	agctcccatt	acacagtgg	tgctcagtgg	2400
gacggccctg	tgggcacctc	cattgctaac	ctcagtgcga	acgatgagga	cacaggagag	2460
aatgcccgc	tcacctacgt	gattcaggac	cccgtgccgc	agttccgc	tgaccccgac	2520
agtggcacca	tgtacaccat	gatggagctg	gactatgaga	accaggtcgc	ctacacgctg	2580
accatcatgg	cccaggacaa	cggcatccc	cagaaatcag	acaccaccac	cctagagatc	2640
ctcatcctcg	atgccaatga	caatgcaccc	cagttcctgt	gggatttcta	ccagggttcc	2700
atctttgagg	atgctccacc	ctcgaccagc	atcctccagg	tctctgccac	ggaccgggac	2760
tcagggtccca	atgggcgtct	gctgtacacc	ttccagggtg	gggacgacgg	cgatggggac	2820
ttctacatcg	agcccacgtc	cgggtgtgatt	cgcaaccagc	gccggctgga	ccgggagaa	2880
gtggccggtg	acaacctttg	ggctctggct	gtggatcggg	gcagtccac	tccccttagc	2940
gcctcggtag	aaatccagg	gaccatcttg	gacattaatg	acaatgcccc	catgtttgag	3000
aaggacgaac	tggagctgtt	tgttgaggag	aacaacccag	tggggtcggt	ggtggcaaa	3060
attcgtgcta	acgaccctga	tgaaggccct	aatgcccaga	tcatgtatca	gattgtggaa	3120
ggggacatgc	ggcatttctt	ccagctggag	ctgctcaacg	gggacctgcg	tgccatggtg	3180
gagctggact	ttgaggtccg	gcgggagtat	gtgctggtgg	tgccaggccac	gtcggtccg	3240
ctggtgagcc	gagccacgg	gcacatcctt	ctcgtggacc	agaatgacaa	ccgcctgtg	3300
ctgcccga	tccagatcct	cttcaacaac	tatgtcacca	acaagtccaa	cagtttcccc	3360
accggcgtag	tcggctgcat	cccggcccat	gaccccgacg	tgctcagacag	cctcaactac	3420
accttcgtgc	agggcaacga	gctgcgcctg	ttgctgctgg	accccgccac	gggcgaactg	3480
cagctcagcc	gcgacctgga	caacaacccg	ccgctggagg	cgctcatgga	ggtgtctgtg	3540

tctgatggca	tccacagcgt	cacggccttc	tgcacctgc	gtgtcaccat	catcacggac	3600
gacatgctga	ccaacagcat	cactgtccgc	ctggagaaca	tgtcccagga	gaagttcctg	3660
tccccgctgc	tggccctctt	cgtggagggg	gtggccgcgc	tgtgtccac	caccaaggac	3720
gacgtcttcg	tcttcaacgt	ccagaacgac	accgacgtca	gctccaacat	cctgaacgtg	3780
accttctcgg	cgctgctgcc	tggcggcgtc	cgcggccagt	tcttcccgtc	ggaggacctg	3840
caggagcaga	tctacctgaa	tgggacgctg	ctgaccacca	tctccacgca	gcgcgtgctg	3900
cccttcgacg	acaacatctg	cctgcgcgag	ccctgcgaga	actacatgaa	gtgcgtgtcc	3960
gttctgcat	tgcacagctc	cgcgcccttc	ctcagctcca	ccaccgtgct	cttccggccc	4020
atccacccca	tcaacggcct	gcgctgccgc	tgcgcgcgcg	gcttcaccgg	cgactactgc	4080
gagacggaga	tgcacctctg	ctactccgac	cgtgcggcgc	ccaacggccg	ctgccgcagc	4140
cgcgagggcg	gctacacgtg	cgagtgtctc	gaggacttca	ctggagagca	ctgtgaggtg	4200
gatgcccgct	caggccgctg	tgccaacggg	gtgtgcaaga	acggggggcac	ctgcgtgaac	4260
ctgctcatcg	ggggcttcca	ctgcgtgtgt	cctcctggcg	agtatgagag	gccctactgt	4320
gaggtgacca	ccaggagctt	cccgccccag	tccttcgtca	ccttccgggg	cctgagacag	4380
cgcttccact	tcaccatctc	cctcacgttt	gccactcagg	aaaggaacgg	cttgcttctc	4440
tacaacggcc	gcttcaatga	gaagcacgac	ttcatcgccc	tggagatcgt	ggacgagcag	4500
gtgcagctca	ccttctctgc	agggcgagaca	acaacgaccg	tggcaccgaa	ggttcccagt	4560
gggtgtgagt	acggggcggtg	gcactctgtg	cagggtgcagt	actacaacaa	gcccaatatt	4620
ggccacctgg	gcttccccca	tgggcccgtcc	ggggaaaaga	tggccgtggg	gacagtggat	4680
gattgtgaca	caaccatggc	tgtgcgcttt	ggaaaggaca	tcgggaacta	cagctgcgct	4740
gcccagggca	ctcagaccgg	ctccaagaag	tccttggtatc	tgaccggccc	totactcctg	4800
gggggtgtcc	ccaacctgcc	agaagacttc	ccagtgcaca	accggcagtt	cgtgggctgc	4860
atgcggaacc	tgtcagtcga	cggcaaaaat	gtggacatgg	ccggattcat	cgccaacaat	4920
ggcaccgggg	aaggctgcgc	tgctcggagg	aacttctgcg	atgggagggc	gtgtcagaat	4980
ggaggcacct	gtgtcaacag	gtggaatatg	tatctgtgtg	agtgtccact	ccgattccgg	5040
gggaagaact	gtgagcaagc	catgcctcac	ccccagctct	tcagcgggtga	gagcgtcgtg	5100
tcctggagtg	acctgaacat	catcatctct	gtgccctggg	acctggggct	catgttccgg	5160
acccggaagg	aggacagcgt	tctgatggag	gccaccagtg	gtggggccac	cagcttccgc	5220
ctccagatcc	tgaacaacta	cctccagttt	gaggtgtccc	acggcccttc	cgatgtggag	5280
tcctgtatgc	tgtccgggtt	gcgggtgacc	gacggggagt	ggcaccacct	gctgatcgag	5340
ctgaagaatg	ttaaggagga	cagtggatg	aagcacctgg	tcaccatgac	cttggactat	5400
gggatggacc	agaacaaggc	agatatcggg	ggcatgcttc	ccgggctgac	ggtaaggagc	5460
gtggtggctg	gaggcgccctc	tgaagacaag	gtctccgtgc	gccgtggatt	ccgaggtgc	5520
atgcagggag	tgaggatggg	ggggacgccc	accaacgtcg	ccaccctgaa	catgaacaac	5580
gcactcaagg	tcagggtgaa	ggacggctgt	gatgtggacg	acccctgtac	ctcgagcccc	5640
tgtcccccca	atagccgctg	ccacgacgcc	tgggaggact	acagctgcgt	ctgtgacaaa	5700
gggtaccttg	gaataaaactg	tgtggatgcc	tgtcacctga	acccctgcga	gaacatgggg	5760
gcctgcgtgc	gctcccccg	ctccccgcag	ggctacgtgt	gcgagtgtgg	gcccagtcac	5820
tacggggccgt	actgtgagaa	caaactcgac	cttccgtgcc	ccagaggtcg	gtgggggaac	5880
cccgctctgtg	gaccctgcca	ctgtgcgctc	agcaaaggct	ttgatcccga	ctgtaataag	5940
accaacggcc	agtgccaatg	caaggagaat	tactacaagc	tcctagccca	ggacacctgt	6000
ctgccctgcg	actgettccc	ccatggctcc	cacagccgca	cttgcgacat	ggccaccggg	6060
cagtgtgcct	gcaagcccg	cgtcatcgcc	cgccagtgca	accgctgcga	caaccctgtt	6120
gcccaggtca	ccacgctcgg	ctgtgaagtg	atctacaatg	gctgtcccaa	agcatttgag	6180
gcccgcactc	ggtggccaca	gaccaagttc	gggcagccgg	ctgcggtgcc	atgccctaag	6240
ggatccgttg	gaaatgcggg	ccgacactgc	agcggggaga	agggctggct	gccccagag	6300
ctctttaact	gtaccaccat	ctccttcgtg	gacctcaggg	ccatgaatga	gaagctgagc	6360
cgcaatgaga	cgcaggtgga	cggcgccagg	gccttcagc	tggtgagggc	gctgcgcagt	6420
gctacacagc	acacgggcac	gctctttggc	aatgacgtgc	gcacggccta	ccagctgctg	6480
ggccacgtcc	ttcagcacga	gagctggcag	cagggtcttc	acctggcagc	cacgcaggac	6540
gcccactttc	acgaggacgt	catccactcg	ggcagcgccc	tcctggcccc	agccaccagg	6600
gcggcgtggg	agcagatcca	gcggagcgag	ggcggcacgg	cacagctgct	ccggcgccctc	6660
gagggctact	tcagcaacgt	ggcacgcaac	gtgcggcgga	cgtacctgcg	gcccttcgtc	6720
atcgtcaccg	ccaacatgat	tcttgcgtgc	gacatctttg	acaagttcaa	ctttacggga	6780
gccagggctc	cgcgattcga	caccatccat	gaagagttcc	ccagggagct	ggagtccctc	6840
gtctccttcc	cagccgactt	cttcagacca	cctgaagaaa	aagaaggccc	cctgctgagg	6900
ccggctggcc	ggaggaccac	cccgcagacc	acgcgcccgg	ggcctggcac	cgagagggag	6960
gccccgatca	gcaggcgagg	gcgacaccct	gatgacgctg	gccagttcgc	cgtgcgtctg	7020
gtcatcattt	accgcaccct	ggggcagctc	ctgcccagac	gctacgaccc	cgaccgtcgc	7080
agcctccggg	tgcttcaccg	gcccacattt	aataccccga	tgggtgagcac	gctgggtgtac	7140
agcgaggggg	ctccgctccc	gagaccctcg	gagagggccg	tcctgggtgga	gttcgcctcg	7200

ctggaggtgg	aggagcgaaac	caagcctgtc	tgcgtgttct	ggaaccactc	cctggccggtt	7260
ggtgggacgg	gaggggtggtc	tgcccggggc	tgcgagctcc	tgtccaggaa	ccggacacat	7320
gtcgcttgc	agtgcagcca	cacagccagc	tttgcggtgc	tcattggatat	ctccaggcgt	7380
gagaacgggg	aggtcctgcc	tctgaagatt	gtcacctatg	ccgctgtgtc	cttgtcactg	7440
gcagccctgc	tggtggcctt	cgctctcctg	agcctgggtcc	gcattgtgcg	ctccaacctg	7500
cacagcattc	acaagcacct	cgccgtggcg	ctcttctctt	ctcagctggg	gttcgtgatt	7560
gggatcaacc	agacggaaaa	cccgtttctg	tgacagtggt	ttgccatcct	cctccactac	7620
atctacatga	gcacctttgc	ctggaccctc	gtggagagcc	tgcatgtcta	ccgcatgctg	7680
accgaggtgc	gcaacatcga	cacggggccc	atgcggttct	actacgtcgt	gggctggggc	7740
atccccggca	ttgtcacagg	actggcggtc	ggcctggacc	cccagggtta	cgggaacccc	7800
gacttctgct	ggctgtcgct	tcaagacacc	ctgatttggg	gctttgcggg	gccccactga	7860
gctgttataa	tcattcaaac	agtcacttct	gtcctatctg	caaaggtttc	ctgccaaaaga	7920
aagcaccatt	attatgggaa	aaaagggatc	gtctccctgc	tgaggaccgc	attcctcctg	7980
ctgctgtcga	tcagcgccac	ctggctgctg	gggctgctgg	ctgtgaaccg	cgatgcactg	8040
agctttcact	acctcttcgc	catcttcagc	ggcttacagg	gccccctcgt	cctccttttc	8100
cactgcgtgc	tcaaccagga	ggtccgggaag	cacctgaagg	gcgtgctcgg	cgggaggaag	8160
ctgcacctgg	aggactccgc	caccaccagg	gccaccctgc	tgacgcgctc	cctcaactgc	8220
aacaccacct	tcggtgacgg	gcctgacatg	ctgcgcacag	acttgggcga	gtccaccgcc	8280
tcgctggaca	gcattcgtcag	ggatgaaggg	atccagaagc	tcggcgtgtc	ctctgggctg	8340
gtgaggggca	gccacggaga	gccagacgcg	tccctcatgc	ccaggagctg	caaggatccc	8400
cctggccacg	attccgactc	agatagcgag	ctgtccctgg	atgagcagag	cagctcttac	8460
gcctcctcac	actcgtcaga	cagcgaggac	gatgggggtg	gagctgagga	aaaatgggac	8520
ccggccaggg	gcgcctgcca	cagcaccccc	aaaggggacg	ctgtggccaa	ccacgttccg	8580
gccggctggc	cgcaccagag	cctggctgag	agtgcacagt	aggacccag	cggcaagccc	8640
cgctgaagg	tggagaccaaa	ggtcagcgtg	gagctgcacc	gcgaggagca	gggcagtcac	8700
cgtggagagt	accccccgga	ccaggagagc	gggggcgag	ccaggcttgc	tagcagccag	8760
ccccagagc	agaggaaagg	catcttgaaa	aataaagtca	cctaccgcgc	gccgctgacg	8820
ctgacggagc	agacgctgaa	gggcgggctc	cggggaagc	tgcccgactg	tgagcagagc	8880
cccacatcct	cgcgcagctc	ttccctgggc	cttggcgccc	ccgactgcgc	catcacagtc	8940
aagagccctg	ggggggagcc	ggggcgtgac	cacctcaacg	gggtggccat	gaatgtgcgc	9000
actgggagcg	cccaggccga	tggtctcgac	tctgagaaac	cgtgaggcaa	gcccgtcacc	9060
ccacacaggc	tgccgcatca	ccctcagacc	ttggagccca	aggggccact	gcccttgaag	9120
tggagtgggc	ccagagtgtg	gcgggtcccca	tggtggcagc	cccccgactg	atcatccaga	9180
cacaaaggct	ttggttctcc	caggagctca	gggcctgtca	gacctgggtga	caagtgcmaa	9240
aggccacagg	catgagggag	gcgtggacca	ctgggcccag	accgctgagt	cctaagactg	9300
cagtcaaaagc	cagaactgag	aggggacccc	agactggggc	cagaggctgg	ccagagttca	9360
ggaacgccgg	gcacagacca	aagaccgcgg	tccagccccg	cccaggcggg	catctcatgg	9420
cagtgcggac	ccgtggctgg	cagcccgggc	agtcctttgc	aaaggcacc	cttgtcttaa	9480
aatcacttctg	ctatgtggga	aagggtggaga	tacttttata	tattttgtatg	ggactctgag	9540
gaggtgcaac	ctgtatatat	attgcattctg	tgtgactttt	gttatcccga	gagatccatg	9600
caatgatctc	ttgtgtctct	ctctgtcaag	attgcacagt	tgtactttaa	ctctggcatgt	9660
gttgacgaaa	ctgggtgccc	agcagatcaa	aggtgggaaa	tacgtcagca	gtggggctaa	9720
aaccaagcgg	ctagaagccc	tacagctgcc	ttcgccaggg	aagtgaggat	gggtgtggggc	9780
ctccccgcgc	gccccctggg	tccccagtg	tcgctgtgtg	tgcgtttgtc	ctctgctgcc	9840
atctgccccg	gctgtgtgaa	ttcaagacag	ggcagtgacg	cactaggcag	gtgtgaggag	9900
ccctgctgag	gtcactgtgg	ggcacggttg	ccacacggct	gtcatttttc	acctggtcat	9960
tctgtgacca	ccacccccctc	ccctcacccg	ctcccagggtg	gcccgggagc	tgccaggtggg	10020
gatggctttg	tccttttgctc	ctgctccccg	tgggacctgg	gaccttaaa	cgttgcagg	10080
tcctgatttg	gacagaggtg	tggggcccctc	caggccgtta	catacctcct	gccaattctc	10140
taactctctg	agactgcgag	gatctccagg	cagggttctc	ccctctggag	tctgaccaat	10200
tacttcattt	tgcttcaaat	ggccaattgt	gcagagggac	aaagccacag	ccacactctt	10260
caacggttac	caaactgttt	ttggaaaattc	acaccaaggt	cgggcccact	gcaggcagct	10320
ggcacagcgt	ggccccgagg	gctgtggaac	gggtcccggg	actgtcagac	atggtttgatt	10380
ttagcgtttc	ctttgttctt	caaatcaggt	gccccaaataa	gtgatcagca	cagctgcttc	10440
caaataggag	aaaccataaaa	ataggatgaa	aatacaagtaa	aatgcaaaga	tgtccacact	10500
gttttaaaact	tgaccctgat	gaaaatgtga	gcactgttag	cagatgccta	tgggagagga	10560
aaagcgtatc	tgaaaatggt	ccaggacagg	aggatgaaat	gagatcccag	agtcctcaca	10620
cctgaatgaa	ttatacatgt	gccttaccag	gtgagtggctc	tttcgaagat	aaaaaactct	10680
agtcccttta	aacgttttgc	cctggcgttt	cctaagtacg	aaaaggtttt	taagtcttctg	10740
aacagtctcc	tttcatgact	ttaacaggat	tctgccccct	gaggtgtaat	ttttttgttc	10800

```

tattttttt cactactcc acagccaaca tcacgaggtg taatttttaa tttgatcaga 10860
actgttacca aaaaacaact gtcagtttta ttgagatggg aaaaatgtaa acctattttt 10920
attacttaag actttatggg agagattaga cactggaggt ttttaacaga acgtgtattt 10980
attaatgttc aaaacactgg aattacaaat gagaagagtc tacaataaat taagattttt 11040
gaatttgtac ttctgcggtg ctggtttttc tccacaaaca cccccgccc tccccatgcc 11100
caggggtggcc gtggaaggga cggtttaagg acgtgcagct gagctgtccg tgtcccacag 11160
tccctcagcc agtggaaact gcccgaactt tttgtccatt ccctagtagg cctgccacag 11220
cctagatggg cagtttttgt ctttcaccaa atttgaggac tttttttttt tgccattatt 11280
tcttcagttt tcttttcttg cactgatctt tctcctctcc ttctgtgact ccagtgactc 11340
agacgttaga cctcttgatg ttttcccact ggtccctgag gctctgttc 11389

```

<210> 238

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014246

<400> 238

```

gggagagatt agacactgga ggtttttaac agaacgtgta tttattaatg ttcaaaacac 60

```

<210> 239

<211> 4372

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014314

<400> 239

```

tagttattaa agttcctatg cagctccgcc tcgctccgg cctcatttcc tcggaaaatc 60
cctgctttcc ccgctcgcca cgccctcctc ctaccgggt ttaaagctag tgaggcacag 120
cctgcgggga acgtagctag ctgcaagcag aggcgggcat gaccaccgag cagcgacgca 180
gcctgcaagc cttccaggat tatatccgga agaccctgga cctacctac atcctgagct 240
acatggcccc ctggttttagg gaggaagagg tgcagtatat tcaggctgag aaaaacaaca 300
agggcccaat ggaggctgcc acactttttc tcaagttcct gttggagctc caggaggaag 360
gctggttccg tggctttttg gatgccctag accatgcagg ttattctgga ctttatgaag 420
ccattgaaag ttgggatttc aaaaaaattg aaaagttgga ggagtataga ttacttttaa 480
aacgtttaca accagaattt aaaaccagaa ttatcccaac cgatatcatt tctgatctgt 540
ctgaatgttt aattaatcag gaatgtgaag aaattctaca gatttgcct actaagggga 600
tgatggcagg tgtagagaaa ttggtggaat gccttctcag atcagacaag gaaaactggc 660
ccaaaacttt gaaacttgct ttggagaaaag aaaggaacaa gttcagtga ctgtggattg 720
tagagaaagg tataaaagat gttgaaacag aagatcttga ggataagatg gaaacttctg 780
acatacagat tttctaccaa gaagatccag aatgccagaa tcttagtgag aattcatgtc 840
caccttcaga agtgtctgat acaaacttgt acagccatt taaaccaaga aattaccaat 900
tagagcttgc tttgcctgct atgaaaggaa aaaacacaat aatatgtgct cctacaggtt 960
gtggaaaaac ctttgtttca ctgcttatat gtgaacatca tcttaaaaaa ttcccacaag 1020
gacaaaaggg gaaagtgtgc ttttttgcga atcagatccc agtgtatgaa cagcagaaat 1080
ctgtattctc aaaatacttt gaaagacatg ggtatagagt tacaggcatt tctggagcaa 1140
cagctgagaa tgtcccagtg gaacagattg ttgagaacaa tgacatcatc attttaactc 1200
cacagattct tgtgaacaac cttaaaaagg gaacgattcc atcactatcc atctttactt 1260
tgatgatatt tgatgaatgc cacaacacta gtaacaaca cccgtacaat atgatcatgt 1320
ttaattatct agatcagaaa cttggaggat cttcaggccc actgcccag gtcattgggc 1380
tgactgcctc gggtggtgtt ggggatgcca aaaacacaga tgaagccttg gattatatct 1440
gcaagctgtg tgcttctctt gatgcgtcag tgatagcaac agtcaaacac aatctggagg 1500
aactggagca agttgtttat aagcccaga agtttttcag gaaagtggaa tcacggatta 1560
gcgacaaatt taaatacatc atagctcagc tgatgagggg cacagagagt ctggcaaaga 1620
gaatctgcaa agacctcgaa aacttatctc aaattcaaaa tagggaattt ggaacacaga 1680
aatatgaaca atggattgtt acagttcaga aagcatgcat ggtgttccag atgccagaca 1740
aagatgaaga gagcaggatt tgtaaagccc tgtttttata cacttcacat ttgcggaaat 1800

```

```

ataatgatgc cctcattatc agtgagcatg cacgaatgaa agatgctctg gattacttga 1860
aagactttctt cagcaatgtc cgagcagcag gattcgatga gattgagcaa gatcttactc 1920
agagattttga agaaaagctg caggaactag aaagtgtttc cagggatccc agcaatgaga 1980
atcctaaact tgaagacctc tgcttcatct tacaagaaga gtaccactta aaccagaga 2040
caataacaat tctctttgtg aaaaccagag cacttgtgga cgctttaaaa aattggattg 2100
aaggaaatcc taaactcagt tttctaaaac ctggcatatt gactggacgt ggcaaaacaa 2160
atcagaacac aggaatgacc ctcccggcac agaagtgtat attggatgca ttcaaagcca 2220
gtggagatca caatattctg attgccacct cagttgctga tgaaggcatt gacattgcac 2280
agtgcaatct tgtcatcctt tatgagtatg tgggcaatgt catcaaatg atccaaacca 2340
gaggcagagg aagagcaaga ggtagcaagt gcttccttct gactagtaat gctggtgtaa 2400
ttgaaaaaga acaataaaac atgtacaaag aaaaaatgat gaatgactct attttacgcc 2460
ttcagacatg ggacgaagca gtatttaggg aaaagattct gcatatacag actcatgaaa 2520
aattcatcag agatagtcaa gaaaaaccaa aacctgtacc tgataaggaa aataaaaaac 2580
tgctctgcag aaagtgcaaa gccttggcat gttacacagc tgacgtaaga gtgatagagg 2640
aatgccatta cactgtgctt ggagatgctt ttaaggaatg ctttgtgagt agaccacatc 2700
ccaagccaaa gcagttttca agttttgaaa atccatgtga agtacaagac atttgagatt ccagttataa 2760
actgcagcca tgactgggga atccatgtga agtacaagac atttgagatt ccagttataa 2820
aaattgaaag ttttgtggtg gaggatattg caactggagt tcagacactg tactcgaagt 2880
ggaaggactt tcattttgag aagataccat ttgatccagc agaaatgtcc aaatgatatc 2940
aggctctcaa tcttcagcta cagggaatga gtaactttga gtggagaaga aacaaacata 3000
gtgggtataa tcatggatcg cttgtacccc tgtgaaaata tattttttta aaatatcttt 3060
agcagtttgt actatattat atatgcaaaag cacaaatgag tgaatcacag cactgagtat 3120
tttgtaggcc aacagagctc atagtacttg gaaaaatta aaaagcctca tttctagcct 3180
tctttttaga gtcaactgcc aacaaacaca cagtaatcac tctgtacaca ctgggataga 3240
tgaatgaatg gaatgttggg aatttttata tccctttgtc tccttaacct actgtaaact 3300
ggcttttgcc cttaacaatc tactgaaatt gttcttttga aggttaccag tgactctggt 3360
tgccaaatcc actgggcact tcttaacctt ctatttgacc tctgcgcatt tggccctggt 3420
gagcaactct cttgaagctc tccctgggct tctctctctt ctagtcttat tctagtcttt 3480
ttttattgag tctcctctt tgctgatccc ttccaaggt tcaatatata tacatgtata 3540
tactgtacat atgtatatgt aactaatata catacatata ggtatgtata tgtaatggtt 3600
atatgtactc atgttcttg tgtagcaacg tgtggtatgg ctacacagag aacatgagaa 3660
cataaagcca tttttatgct tactactaaa agctgtccac tgtagagttg ctgtatgtag 3720
caatgtgtat ccactctaca gtggtcagct tttagtagag agcataaaaa tgataaaata 3780
cttcttgaaa acttagttta ctatacatct tgccctatta atatgttctc ttaacgtgtg 3840
ccattgttct ctttgacctt ttctctataa tgatgttgat gttcaacacc tggactgaat 3900
gtctgttctc agatcccttg gatgttacag atgaggcagt ctgactgtcc tttctacttg 3960
aaagattaga atatgtatcc aaatggcatt cacgtgtcac ttagcaagggt ttgctgatgc 4020
ttcaaagagc ttagtttgcg gtttcttgga cgtggaaaca agtatctgag ttccctggag 4080
atcaacggga tgaggtgtta cagctgcctc cctcttcatg caatctggtg agcagtggtg 4140
caggcgggga gccagagaaa cttgccagtt atataaactc tctttggctt ttcttcatct 4200
gtaaaacaag gataatactg aactgtaagg gttagtggag agtttttaat taaaagaatg 4260
tgtgaaaagt acatgacaca gtagtgtctt gataatagtt actagtagta gtattcttac 4320
taagacccaa tacaatgga ttattttaa caaaaaaaaa aaaaaaaaaa aa 4372

```

<210> 240

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014314

<400> 240

agttcagaca ctgtactcga agtggaaagga ctttcatttt gagaagatac catttgatcc 60

<210> 241

<211> 1647

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014321

<400> 241

```

gcgcgcggggt ttcgttgacc cgcggcggttc acgggaattg ttgcgttttag tgccggcgcc 60
atgggggtcgg agctgacggg gcgcctagcc ccgcgcctgg gcctcgccga gcccgacatg 120
ctgaggaaag cagaggagta cttgcgcctg tcccgggtga agtggtgctgg cctctccgca 180
cgcaccacgg agaccagcag tgcagtcacg tgccctggacc ttgcagcttc ctggatgaag 240
tgcccccttg acagggctta tttaattaaa ctttctgggt tgaacaagga gacatatcag 300
agctgtctta aatcttttga gtgtttactg ggcctgaatt caaatattgg aataagagac 360
ctagctgtac agtttagctg tatagaagca gtgaacatgg cttcaaagat actaaaaagc 420
tatgagtcga gtcttcccca gacacagcaa gtggatcttg acttatccag gccacttttc 480
acttctgctg cactgctttc agcatgcaag attctaaagc tgaaagtggg taaaaacaaa 540
atggtagcca catccgggtg aaaaaaagc atatttgatc gactgtgtaa acaactagag 600
aagattggac agcaggctga cagagaacct ggagatgtag ctactccacc acggaagaga 660
aagaagatag tggttgaagc cccagcaaag gaaatggaga aggtagagga gatgccacat 720
aaaccacaga aagatgaaga tctgacacag gattatgaag aatggaaaag aaaaattttg 780
gaaaatgctg ccagtgctca aaaggctaca gcagagtgat ttcagcttcc aaactggtat 840
acattccaaa ctgatatgac attgccatct ccaggaagac ttgacggctt tgggattttg 900
tttaaaacttt tataataagg atcctaagac tgttgccttt aaatagcaaa gcagcctacc 960
tggaggctaa gtctgggcag tgggctggcc cctgggtgtg gcattagacc agccacagtg 1020
cctgattggg atagccttat gtgctttcct acaaaatgga attggaggcc gggcgagtg 1080
gtcacgcct gtaatcccag cactttggga ggccaagggt ggtggatcac ctgaggtcag 1140
gagctcgaga ccagcctggc caacatgggt aaaccccatc tctactaaaa atacaaaaat 1200
tagccagggt tgatgggtga tgcctgtaat cccagctcct cagtaggctg agacaggagc 1260
atcacttgaa cgtgggaggg agaggttgca gtgagccgag attgcaccac cgcactccag 1320
cctgggtgac agagcgagac ttatctcata aataaataga tagatactcc agcctgggtg 1380
acagagcgag acttatagat agatagatag atagatggat agatagatag atagatagat 1440
agatagataa acggaattgg agccattttg ctttaagtga atggcagtc cttgtcttat 1500
tcagaatata aaattcagtc tgaatggcat cttacagatt ttacttcaat tttgtgtac 1560
ggatattttt atttgactaa atcaatatat tgtacgcct aagttaataa atgttattta 1620
tatatgcaaa aaaaaaaaaa aaaaaaa 1647

```

<210> 242

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014321

<400> 242

```

tgctttaagt gaatggcagt cccttgtctt attcagaata taaaattcag tctgaatggc 60

```

<210> 243

<211> 1455

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014364

<400> 243

```

ggcggtccgc acgcacctcg gtaacatcac agcagggtcca ggccaatgat aaccttataa 60
gaggccatgt cgaagcgcgga catcgctcct accaatgtca ccgttgtcca gttgctgcga 120
cagccgtgcc cggtgaccag agcaccgccc ccacctgagc ctaaggctga agtagagccc 180
cagccacaac cagagcccac accagtcagg gaggaataa agccaccacc gccaccactg 240
cctctcacc ccgtactcc tctctctaag atggtgtctg tggcccgga gctgactgtg 300
ggcatcaatg gatttggacg catcggtcgc ctggtcctgc gcgcctgcat ggagaagggg 360
gttaaggtgg tggctgtgaa tgatccattc attgaccggg aatacatggg gtacatgttt 420
aagtatgact ccaccacgg ccgatacaag ggaagtgtgg aattcaggaa tggacaactg 480
gtcgtggaca accatgagat ctctgtctac cagtgcgaag agcccaaaca gatcccctgg 540

```

```

agggtctgtcg ggagccccta cgtgggtggag tccacaggcg tgtacctctc catacaggca 600
gcttcggacc acatctctgc aggtgctcaa cgtgtggtca tctccgcgcc ctcaccggat 660
gcaccaatgt tcgtcatggg tgccaatgaa aatgactata accctggctc catgaacatt 720
gtgagcaacg cgtcctgcac caccaactgt ttggctcccc tcgccaaagt catccacgag 780
cgatttggga tcgtggaagg gttgatgacc acagtccatt cctacacggc caccagaag 840
acagtggacg ggccatcaag gaaggcctgg cgagatgggc ggggtgcca ccagaacatc 900
atcccagcct ccaactggggc tgcgaaagct gtgaccaaag tcatcccaga gctcaaaggg 960
aagctgacag ggatggcggt cccgggtacca accccggatg tgtctgtcgt ggacctgacc 1020
tgccgcctcg cccagcctgc cccctactca gccatcaagg aggtctgtaa agcagcagcc 1080
aaggggcccc tggctggcat ccttgccctac accgaggatg aggtcgtctc tacggacttc 1140
ctcggtgata cccactcgtc catcttcgat gctaaggccg gcattgcgct caatgacaat 1200
ttcgtgaagc tcatttcatt gtacgacaac gaatatggct acagtcaccg ggtggctgac 1260
ctcctccgct acatgttcag ccgagacaag tgaaacggga aggtcccttc tttccttccc 1320
aggggcccgg gccggaacat gtgcctcccg ttccagcatt tggtgcccg ggggaggaag 1380
gacaccggg gcggggcgccc cagcccgatg ggtccatggt gaaataaaaa acagtgtctg 1440
aaaaaaaaa aaaaa 1455

```

<210> 244

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014364

<400> 244

cgctcaatga caatttcgtg aagctcattt catggtacga caacgaatat ggctacagtc 60

<210> 245

<211> 935

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014462

<400> 245

```

gaagtgggta agggtaatat ggaggagctt cgggcaggcc cgggcgggctg aaagccgggg 60
cagaagtgct ggtctcggtc gggattccgg gcttgggtccc accgaggcgg cgactgcggg 120
aggagggaag aggttttggg cgcgctggcc tcccgcgcgt gtgcattgca gcattatttc 180
agttcaaaat gaactatatg octggcaccg ccagcctcat cgaggacatt gacaaaaagc 240
acttggttct gcttcgagat ggaaggacac ttataggctt tttaagaagc attgatcaat 300
ttgcaaaactt agtgctacat cagactgtgg agcgtattca tgtgggcaaa aaatacgggtg 360
atattcctcg agggattttt gtggtcagag gagaaaatgt ggtcctacta ggagaaatag 420
acttggaaaa ggagagtgac acaccctcc agcaagtatc cattgaagaa attctagaag 480
aacaagggtt ggaacagcag accaagctgg aagcagagaa gttgaaagtg caggccctga 540
aggaccgagg tctttccatt cctcgagcag atactccttg tgagtactaa tcttttgccc 600
agaggctggt ggctcttgaa gagtaggggc tgtoactgag tgaaagtgac atcctggcca 660
cctcacgcat ttgatcacag actgtagagt tttgaaaagt cactttttatt ttttaattatt 720
ttacatatgc aacatgaaga aatcgtgtag gtgggttttt tttttaataa caaaatcact 780
gtttaaagaa acagtggcat agactccttc acacatcact gtggcaccag caactacttc 840
tttatattgt tcttcatatc ccaaattaga gtttacaggg acagtcttca tttactttga 900
aataaaatat gaatctcaaa aaaaaaaaaa aaaaa 935

```

<210> 246

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014462

<400> 246
ttaataacaa aatcactgtt taaagaaaca gtggcataga ctccctcaca catcactgtg 60

<210> 247
<211> 890
<212> DNA
<213> Homo sapiens

<300>
<308> NM_014501

<400> 247
ggcggaccga agaacgcagg aagggggccg gggggacccg cccccggccg gccgcagcca 60
tgaactccaa cgtggagaac ctacccccgc acatcatccg cctgggtgtac aaggagggtga 120
cgacactgac cgcagaccca cccgatggca tcaaggtctt tcccaacgag gaggacctca 180
ccgacctcca ggtcaccatc gagggccctg aggggacccc atatgctgga ggtctgttcc 240
gcatgaaaact cctgctgggg aaggacttcc ctgcctcccc acccaagggc tacttcctga 300
ccaagatctt ccacccgaac gtgggcgcca atggcgagat ctgctgcaac gtgctcaaga 360
gggactggac ggctgagctg ggcattccgac acgtactgct gaccatcaag tgcctgctga 420
tccacctaa ccccgagtct gcactcaacg aggaggcggg ccgcctgctc ttggagaact 480
acgaggagta tgcggctcgg gcccgctctg tcacagagat ccacgggggc gccggcgggc 540
ccagcggcag ggccgaagcc ggtcggggcc tggccagtgg cactgaagct tcctccaccg 600
accctggggc cccagggggc ccgggagggg ctgagggtcc catggccaag aagcatgctg 660
gcgagcgcga taagaagctg gcggccaaga aaaagacgga caagaagcgg gcgctgcggg 720
cgctgcggcg gctgtagtgg gctctcttcc tccttccacc gtgaccccaa cctctcctgt 780
cccctccctc caactctgtc tctaagttat ttaaattatg gctggggctc gggagggtac 840
agggggcact gggacctgga tttgtttttc taaataaagt tggaaaagca 890

<210> 248
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_014501

<400> 248
acacgtactg ctgaccatca agtgccctgt gatccaccct aaccccgagt ctgcactcaa 60

<210> 249
<211> 1182
<212> DNA

<213> Homo sapiens

<300>
<308> NM_016095

<400> 249
gcggccggcg gcgtctcctc ccgggacgct gaggggcccg aggagaccgt gaggtctctg 60
cctgcagctc gcgcgcgaat ggacgctgcc gaggtogaat tcctcgccga gaaggagctg 120
gttaccatta tccccactt cagtctggac aagatctacc tcctcggggg ggacctggg 180
ccttttaacc ctggtttacc cgtggaagtg cccctgtggc tggcgattaa cctgaaacaa 240
agacagaaat gtgcctgct cctccagag tggatggatg tagaaaagtt ggagaagatg 300
agggatcatg aacgaaagga agaaactttt accccaatgc ccagccctta ctacatggaa 360
cttacgaagc tcctgttaaa tcattgcttca gacaacatcc cgaaggcaga cgaaatccgg 420
accctgggtca aggatatgtg ggacactcgt atagccaaac tccgagtgtc tgctgacagc 480
tttgtgagac agcaggaggc acatgccaaag ctggataact tgaccttgat ggagatcaac 540
accagcggga ctttctctac acaagcgctc aaccacatgt acaaactccg caccgaacctc 600

```

cagcctcttg agagtactca gtctcaggac ttctagagaa aggcctgggtg caggcgggctt 660
gctgggggat gtgagcgctc aggatgtgat gaggtactcg tggttctgga gctctagaaa 720
cacttctgat gcatgaaaaa tgtgtgatgg tgcaaggaat ggattcagga tgttggtgga 780
gaaacaagtt tgtgattagt ccttaaaact tagctccctg ggacattctt caattccaca 840
tctgtttcta gaaaccagcc ctttttcccc ccacttttga gaaataaaaa agccttaggt 900
aaataagtca ttctccctag cagagccact tgggtctcct gcatggaagc cgtcacactt 960
gggcagggtg tcagtgaactg gtaggtgtag atacagcagg agtggccatg tgggccacgg 1020
ctttttaccc cttcttgatc ctgatttctt gggctgaatt tagactctct cacagagggtg 1080
gctcacagag aaggatggca gatgggtgcag ccaacaatgc tgaccgggtgc ttatcctcta 1140
agccctgatc cacaataaaa atggacccaa ctcaaaaaaa aa 1182

```

<210> 250

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016095

<400> 250

```

atggattcag gatgttggtg gagaaacaag tttgtgatta gtccttaaaa cttagctccc 60

```

<210> 251

<211> 704

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016185

<400> 251

```

tgcagcgggtg gtcggctggt ggggtgtggag tttcccagcg cccctcgggt ccgacccttt 60
gagcgttctg ctccggcgcc agcctacctc gctcctcggc gccatgacca caaccaccac 120
cttcaaggga gtcgacccca acagcaggaa tagctcccga gttttgcggc ctccagggtg 180
tggatccaat ttttcattag gttttgatga accaacagaa caacctgtga ggaagaacaa 240
aatggcctct aatatctttg ggacacctga agaaaatcaa gcttcttggg ccaagtcagc 300
aggtgccaag tctagtgggt gcagggaaga cttggagtca tctggactgc agagaaggaa 360
ctcctctgaa gcaagctccg gagacttctt agatctgaag ggagaagggt atattcatga 420
aaatgtggac acagacttgc caggcagcct ggggcagagt gaagagaagc ccgtgcctgc 480
tgcgccctgtg ccagcccggt tggcccggtg cccagtgcc tccagaagaa atccccctgg 540
cggcaagtcc agcctcgtct tgggttagct ctgactgtcc tgaacgctgt cgttctgtct 600
gtttcctcca tgcttgagaa ctgcacaact tgagcctgac tgtacatctt cttggatttg 660
tttcattaaa aagaagcact ttatgtaaaa aaaaaaaaaa aaaa 704

```

<210> 252

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016185

<400> 252

```

tgaaccaaca gaacaacctg tgaggaagaa caaatggcc tctaatatct ttgggacacc 60

```

<210> 253

<211> 2268

<212> DNA

<213> Homo sapiens

<220>

<221> Modified_base
 <222> 1 ... 2268
 <223> n = a,c,g, or t

<300>
 <308> NM_016359

<400> 253
 gggatttgaa ccncgctgac gaagtttggg gatccatctt ccgagtatcg ccgggatttc 60
 gaatcgcat gatcatcccc tctctagagg agctggactc cctcaagtac agtgacctgc 120
 agaacttagc caagagtctg ggtctccggg ccaacctgag ggcaaccaag ttgttaaaag 180
 ccttgaaagg ctacattaaa catgaggcaa gaaaaggaaa tgagaatcag gatgaaagtc 240
 aaacttctgc atcctcttgt gatgagactg agatacagat cagcaaccag gaagaagctg 300
 agagacagcc acttgccat gtcacccaaa caaggagaaag gtgcaagact gtccgtgtgg 360
 accctgactc acagcagaat cattcagaga taaaaataag taatcccact gaattccaga 420
 atcatgaaaa gcaggaaagc caggatctca gagctactgc aaaagttcct tctccaccag 480
 acgagcacca agaagctgag aatgctgttt cctcaggtaa cagagattca aaggtaacct 540
 cagaaggaaa gaaatctctc tacacagatg agtcatccaa acctggaaaa aataaaagaa 600
 ctgcaatcac tactccaaac tttaagaagc ttcatgaagc tcattttaag gaaatggagt 660
 ccattgatca atatatgtg agaaaaagaa acattttgaa gaacacaatt ccatgaatga 720
 actgaagcag cagcccatca ataagggagg ggtcaggact ccagtacctc caagaggaag 780
 actctctgtg gcttctactc ccacagcca acgacgctcg caaggccggg cttgtggccc 840
 tgcaagtcag agtaccttgg gtctgaaggg gtcactcaag cgctctgcta tctctgcagc 900
 taaaacgggt gtcaggtttt cagctgtctac taaagataat gagcataagc gttcactgac 960
 caagactcca gccagaaagt ctgcacatgt gaccgtgtct gggggcacc caaaaggcga 1020
 gggtgtgctt gggacacaca aattaaagac catcacgggg aattctgctg ctgttattac 1080
 ccatttcaag ttgacaactg aggcaacgca gactccagtc tccaataaga aaccagtgtt 1140
 tgatcttaaa gcaagtttgt ctctgtccct caactatgaa ccacacaaag gaaagctaaa 1200
 accatggggg caatctaaag aaaataatta tctaaatcaa catgtcaaca gaattaaact 1260
 ctacaagaaa acttacaac aaccccatct ccagacaaag gaagagcaac ggaagaaacg 1320
 cgagcaagaa cgaaaggaga agaaagcaaa ggttttggga atgcgaagg gcctcatatt 1380
 ggctgaagat taataatttt ttaatatctt gtaaattatc ctgtattctc aacttttttc 1440
 cttttgtaaa tttttttttt ttgtgtgtca tccccacttt agtcacgaga tctttttctg 1500
 ctaactgttc atagtctgtg tagtggtccat gggttcttca tgtgctatga tctctgaaaa 1560
 gacgttatca ccttaaagct caaattcttt gggatgggtt ttacttaagt ccattaacaa 1620
 ttcaggtttc taacgagacc catcctaata ttctgtttct agatttttaa tgtcaagttc 1680
 ccaagttccc cctgtgggtt ctaatatata cagaactgca gtcttctgct agccaatagc 1740
 atttacctga tggcagctag ttatgcaagc ttcaggagaa tttgaacaat aacaagaata 1800
 gggtaagctg ggatagaaag gccacctctt cactctctat agaatatagt aacctttatg 1860
 aaacggggcc atatatgttg gttatgacat caatatatta cctaggtgaa attgttttagg 1920
 cttatgtacc ttctgttcaa tatcctcatg taattgacct ctgtcactca ctatattcac 1980
 aaaaaataaa ctctacaact cattctaaca ttgcttactt aaaagctaca tagccctatc 2040
 gaaatcgag gattaatgct ttaatgcttt tagagacagg gtctcactgt gttgcccagg 2100
 ctggctcaca actccaccaa atgtacttct tattcatttt atggaaaaga ctaggctttg 2160
 cttagtatca tgtccatgtt tcttccacct cagtggagct tctgagtttt atactgctca 2220
 agatcgtcat aaataaaatt ttttctcatt gtcaaaaaaa aaaaaaaa 2268

<210> 254
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_016359

<400> 254
 acattgctta cttaaaagct acatagccct atcgaaatgc gaggattaat gctttaatgc 60

<210> 255
 <211> 1590
 <212> DNA

<213> Homo sapiens

<300>

<308> NM_016816

<400> 255

```

gagggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccaccagcc 60
aaatctcttg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120
atcgaccatg ccattgacat catctgtggg ttcttgaagg aaagggtgctt ccgaggtagc 180
tctaccctg tgtgtgtgtc caagggtggtt aagggtggct cctcaggcaa gggcaccacc 240
ctcagaggcc gatctgacgc tgacctgggt gtcttctctc gtcctctcac cacttttcag 300
gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360
caaagagaga gagcactttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420
cgtgcgctca gcttcgtact gagttcgctc cagctcgggg agggggtgga gttcgatgtg 480
ctgcctgcct ttgatgccct gggtcagttg actggcagct ataaacctaa ccccaaatc 540
tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600
ttcacagaac tacagagaga ctctctgaag cagcgcccca ccaagctcaa gacctcatc 660
cgctagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720
tatgccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780
aacacagccc aaggatttct gacgggtctt gaattagtca taaactacca gcaactctgc 840
atctactgga caaagtatta tgactttaaa aacccatta ttgaaaagta cctgagaagg 900
cagctcacga aaccaggcc tgtgatcctg gaccggcgag accctacagg aaacttgggt 960
gggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020
ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct ggctgaaagc 1080
aacagtacag acgatgagac cgacgatccc aggacgtatc agaaatatgg ttacattgga 1140
acacatgagt accctcattt ctctcataga ccacgacgc tccaggcagc atccacccca 1200
caggcagaag aggactggac ctgcaccatc ctctgaatgc cagtgcactc tgggggaaag 1260
ggctccagtg ttatctggac cagttccttc attttcaggt gggactcttg atccagagaa 1320
gacaaagctc ctcatgagc tgggtgtataa tccaagacag aacccaagtc tctgactcc 1380
tggccttcta tgccctctat cctatcatag ataacattct ccacagctc acttcattcc 1440
acctattctc tgaaaatatt ccctgagaga gaacagagag atttagataa gagaatgaaa 1500
ttccagcctt gactttcttc tgtgcacctg atgggagggt aatgtctaat gtattatcaa 1560
taacaataaa aataaagcaa ataccaaaaa 1590

```

<210> 256

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016816

<400> 256

```

cgatcccagg acgtatcaga aatatgggta cattggaaca catgagtacc ctcatctctc 60

```

<210> 257

<211> 2905

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016817

<400> 257

```

cggcagccag ctgagagcaa tgggaaatgg ggagtcacag ctgtcctcgg tgcttctc 60
gaagctgggt tggtttatcc aggaatacct gaagccctac gaagaatgtc agacactgat 120
cgacgagatg gtgaacacca tctgtgacgt ctgcagggaac cccgaacagt tccccctgg 180
gcaggagtg gccataggtg gtccttatgg acggaaaaca gtcttaagag gcaactccga 240
tggtaccctt gtcttttctc tcagtgaactt aaaacaattc caggatcaga agagaagcca 300
acgtgacatc ctcgataaaa ctgggggataa gctgaagttc tgtctgttca cgaagtgggt 360

```

```

gaaaaacaat ttcgagatcc agaagtcctt tgatgggtcc accatccagg tgttcacaaa 420
aaatcacaga atctcttttg aggtgctggc cgccttcaac gctctgagct taaatgataa 480
tcccagcccc tggatctatc gagagctcaa aagatccttg gataagacaa atgccagttc 540
tgggtgagttt gcagttctgt tcaactgaact ccagcagaag ttttttgaca accgtcctgg 600
aaaactaaag gatttgatcc tcttgataaa gcaactggcat caacagtgcc agaaaaaaat 660
caaggattta cctctgctgt ctccgtatgc cctggagctg cttacggtgt atgcctggga 720
acaggggtgc agaaaagaca actttgacat tgctgaaggc gtcagaacgg ttctggagct 780
gatcaaatgc caggagaagc tgtgtatcta ttggatggtc aactacaact ttgaagatga 840
gaccatcagg aacatcctgc tgcaccagct ccaatcagcg aggccagtaa tcttggatcc 900
agttgaccca accaataatg tgagtggaga taaaatatgc tggcaatggc tgaaaaaaga 960
agctcaaacc tggttgactt ctcccaacct ggataatgag ttacctgcac catcttggaa 1020
tgtctgcctt gcaccactct tcacgacccc aggccacctt ctggataagt tcatcaagga 1080
gtttctccag cccaacaaat gcttcctaga gcagattgac agtgctgtta acatcatccg 1140

```

```

tacattcctt aaagaaaact gcttccgaca atcaacagcc aagatccaga ttgtccgggg 1200
aggatcaacc gccaaaggca cagctctgaa gactggctct gatgccgatc tcgtcgtgtt 1260
ccataactca cttaaaagct acacctccca aaaaaacgag cggcacaaaa tcgtcaagga 1320
aatccatgaa cagctgaaag ccttttggag ggagaaggag gaggagcttg aagtcagctt 1380
tgagcctccc aagtgggaag ctcccagggt gctgagcttc tctctgaaat ccaaagtcct 1440
caacgaaagt gtcagctttg atgtgcttcc tgcctttaat gcaactgggtc agctgagttc 1500
tggctccaca cccagccccg aggtttatgc agggctcatt gatctgtata aatcctcgga 1560
cctcccgga ggagagtttt ctacctgttt cagagtcctg cagcgaaact tcattcgctc 1620
ccggcccacc aaactaaagg atttaattcg cctgggtgaag cactgggtaca aagagtgtga 1680
aaggaaaactg aagccaaagg ggtctttgcc cccaaagtat gccttggagc tgctcaccat 1740
ctatgcctgg gagcagggga gtggagtgc ggattttgac actgcagaag gtttccggac 1800
agtctggag ctggtcacac aatatcagca gctcggcatc ttctggaagg tcaattacaa 1860
ctttgaagat gagaccgtga ggaagtttct actgagccag ttgcagaaaa ccaggcctgt 1920
gatcttggac ccaggcgaac ccacaggtga cgtgggtgga ggggaccgtt ggtgttggca 1980
tcttctggac aaagaagcaa aggttaggtt atoctctccc tgcctcaagg atgggactgg 2040
aaaccaata ccaccttga aagtgcgcac aatgcagaca ccaggaagtt gtggagctag 2100
gatccatcct attgtcaatg agatgttctc atccagaagc catagaatcc tgaataataa 2160
ttctaaaaga aacttctgga gatcatctgg caatcgcttt taaagactcg gctcaccgtg 2220
agaaaagatc actcacatcc attcttccct tgatgggtccc tattcctcct tcccttgctt 2280
tcttggactt cttgaaatca atcaagactg caaaccttt cataaagctg ccttgctgaa 2340
ctctctctg caggagccct gcttaaaata gttgatgtca tcactttatg tgcatttat 2400
ttctgtcaac ttgtattttt ttttcttgta tttttccaat tagctcctcc tttttccttc 2460
cagtctaaaa aaggaatcct ctgtgtcttc aaagcaaagc tctttacttt ccccttgggt 2520
ctcataactc tgtgatcttg ctctcggtgc ttccaactca tccacgtcct gtctgtttcc 2580
tctgtataca aaacctttc tgcccctgct gacacagaca tctctatgc cagcagccag 2640
gccaacctt tcattagaac ttcaagctct ccaaaggctc agattataac tgttgtcata 2700
tttatatgag gctgttgtct tttcctctct agcctgcctt tatccccca cccaggagta 2760
tcctcttgcc aaagcaaaag actttttcct tggcttttagc cttaaagata cttgaaggtc 2820
taggtgcttt aacctcacat accctcaact aaacttttat cactgttgca tataccagtt 2880
gtgatacaat aaagaatgta tctgg 2905

```

<210> 258

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016817

<400> 258

aaggcttagg tgctttaacc tcacataccc tcaacttaaac ttttatcact gttgcatata 60

<210> 259

<211> 2054

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017414

<400> 259

```

gggaagctcg ggccggcagg gtttccccgc acgctggcgc ccagctccccg gcgcggagggc 60
cgctgtaagt ttcgctttcc attcagtggg aaacgaaagc tgggcgggggt gccacgagcgc 120
cggggccaga ccaaggcggg cccggagcgc aacttcggtc ccagctcgggt ccccggtcca 180
gtcccagact ggaactcagc agcggaggct ggacgcttgc atggcgcttg agagattcca 240
tcgtgcctgg ctacataag cgcttcctgg aagtgaagtc gtgctgtcct gaacgcgggc 300
caggcagctg cggcctgggg gttttggagt gatcacgaat gagcaaggcg tttgggctcc 360
tgaggcaaat ctgtcagtc atcctggctg agtcctcgca gtccccggca gatcttgaag 420
aaaagaagga agaagacagc aacatgaaga gagagcagcc cagagagcgt cccagggcct 480
gggactaccc tcatggcctg gttgggtttac acaacattgg acagacctgc tgccttaact 540
ccttgattca ggtgttcgta atgaatgtgg acttcaccag gatattgaag aggatcacgg 600
tgcccagggg agctgacgag cagaggagaa gcgtcccttt ccagatgctt ctgctgctgg 660
agaagatgca ggacagccgg cagaaagcag tgcggccctt ggagctggcc tactgcctgc 720
agaagtgcaa cgtgcccttg tttgtccaac atgatgtgc ccaactgtac ctcaaactct 780
ggaacctgat taaggaccag atcactgatg tgcacttggg ggagagactg caggccctgt 840
atacgatccg ggtgaaggac tccttgattt gcgttgactg tgccatggag agtagcagaa 900
acagcagcat gctcaccctc ccactttctc tttttgatgt ggactcaaag cccctgaaga 960
cactggagga cgccctgcac tgcttcttcc agcccagggg gttatcaagc aaaagcaagt 1020
gcttctgtga gaactgtggg aagaagaccc gtgggaaaca ggtcttgaag ctgacctatt 1080
tgccccagac cctgacaatc cacctcatgc gattctccat caggaattca cagacgagaa 1140
agatctgcca ctccctgtac ttccccaga gcttggattt cagccagatc cttccaatga 1200
agcgagagtc ttgtgatgct gaggagcagt ctggagggca gtatgagctt tttgctgtga 1260
ttgcgcacgt ggggaatggc gactccggtc attactgtgt ctacatccgg aatgctgtgg 1320
atggaaaatg gttctgcttc aatgactcca atatttgctt ggtgtcctgg gaagacatcc 1380
agtgtacctc cggaaatcct aactaccact ggcaggaaac tgcatactt ctgggtttaca 1440
tgaagatgga gtgctaattg aaatgcccaa aaccttcaga gattgacacg ctgtcatttt 1500
ccatttcctg tcctggatct acggagtctt ctaagagatt ttgcaatgag gagaagcatt 1560
gttttcaaac tatataactg agccttattt ataattaggg atattatcaa aatatgtaac 1620
catgaggccc ctccaggtcct gatcagtcag aatggatgct ttcaccagca gaccggcca 1680
tgtggctgct cggctcctggg tgctcgctgc tgtgcaagac attagccctt tagttatgag 1740
cctgtgggaa cttcaggggt tcccagtggt gagagcagtg gcagtgggag gcatctgggg 1800
gccaaaggct agtggcaggg ggtatttcag tattatacaa ctgctgtgac cagacttgta 1860
tactggctga atatcagtg tggtttgta ttttcacttt gagaaccaac attaattcca 1920
tatgaatcaa gtgttttgta actgctattc atttattcag caaataattt ttgatcatct 1980
cttctccata agatagtgtg ataaacacag tcatgaataa agttattttc cacaaaaaaa 2040
aaaaaaaaa aaaa 2054

```

<210> 260

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017414

<400> 260

```

tgagcatctc ttctccataa gatagtgtga taaacacggg catgaataaa gttattttcc 60

```

<210> 261

<211> 3638

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017523

<400> 261

```

ggtagatgcg gctgtgacag cagcaaagaa tgacggccaa gggcgacagc aggggctggc 60

```

catgctgtaa	aggggcttct	tgggaggggc	cagcctcagg	aatcaagggg	aactcctgag	120
ccgagaattc	tgaagatctc	ctccctccct	gaagctgtgg	gctgggccat	cggaaaactt	180
tcagttttgt	ttccttgcc	gcaagaaacg	aaactcaacc	gaaagcctgc	agagagcaga	240
acatggaagg	agacttctcg	gtgtgcagga	actgtaaaag	acatgtagtc	tctgccaaact	300
tcaccctcca	tgaggcttac	tgccctgcgg	tcctggctct	gtgtccggag	tgtgaggagc	360
ctgtccccaa	ggaaaccatg	gaggagcact	gcaagcttga	gcaccagcag	gttgggtgta	420
cgatgtgtca	gcagagcatg	cagaagtcct	cgctggagtt	tcataaggcc	aatgagtgcc	480
aggagcgccc	tgttgagtgt	aagttctgca	aactggacat	gcagctcagc	aagctggagc	540
tccacgagtc	ctactgtggc	agccggacag	agctctgcca	aggctgtggc	cagttcatca	600
tgcaccgcat	gctcgccag	cacagagatg	tctgtcgag	tgaacaggcc	cagctcgga	660
aaggggaaag	aatttcagct	cctgaaagg	aaatctactg	tcattattgc	aaccaaata	720
ttccagaaaa	taagtatttc	caccatatgg	gtaaatgttg	tccagactca	gagtttaaga	780
aacactttcc	tgttggaat	ccagaaatcc	ttccttcctc	tcttccaagt	caagctgctg	840
aaaatcaaac	ttccacgatg	gagaaagatg	ttcgtccaaa	gacaagaagt	ataaacagat	900
ttcctcttca	ttctgaaagt	tcatacaaga	aagcaccag	aagcaaaaac	aaaaccttgg	960
atccactttt	gatgtcagag	cccaagccca	ggaccagctc	ccctagagga	gataaagcag	1020
cctatgacat	tctgaggaga	tgttctcagt	gtggcatcct	gcttcccttg	ccgataccta	1080
atcaacatca	ggagaaatgc	cggtgggttag	cttcatcaaa	aggaaaaaca	gtgagaaatt	1140
tcagctagat	ttggaaaagg	aaaggactta	caaattcaaa	agatttcaact	tttaacactg	1200
gcattcctgc	ctacttgctg	tgggtggtctt	gtgaaagggtg	atgggtttta	ttcgttgggc	1260
tttaaaagaa	aaggtttggc	agaactaaaa	acaaaactca	cgtatcatct	caatagatac	1320
agaaaaggct	tttgataaaa	ttcaacttga	cttcatgtta	aaaacctca	acaaaccagg	1380
cgctgaagga	acatacctca	aaataataag	agccatctat	gacaaaacca	cagccaacat	1440
catactgaat	gagcaaaagc	tggagcatta	ctcttgagaa	gtagaacaag	gcacttcagt	1500
cctattcaac	atagtactgg	aagtctctgc	cacagcaatc	aggcaagaga	aagaaataaa	1560
aggcaaccaa	aaagaaagga	agtcgaagta	tctctgtttg	cagacgatat	gattctatat	1620
ctagaaaacc	ccatgatctt	ggcccaaaag	ctcctagatc	tgataaaaca	cttcagctaa	1680
ctttcaggag	acaaaatcaa	tatacaaaat	atggtagcat	ttttatacac	caacgacatc	1740
caagctgaga	gccaaatcaa	gaatgcaatc	ctattcacaa	ttgccacaaa	aagaataaaa	1800
tacctagga	tacagctaac	caggagatg	aaagatctct	acaacaaaaa	ttacaaaaa	1860
ctgctgaaag	aaatcagaga	tgacacaaat	ggaaaaacat	tccatactta	tggataggaa	1920
gaatcaatat	tgttaaaatg	gccatactac	ccaaagcaat	ttatagattc	aatgctattc	1980
ctatcaaact	accaataaca	ttcttcacag	aatcagaaaa	aaaaagcatt	aaaatttatt	2040
tgaaccacaa	aaagagccca	aaaagccaaa	gcaatcctaa	gcaaaaagaa	caaagctgga	2100
ggcatcgcat	tacccaactt	caaactatac	tacagggcta	cagtaaccac	aactgcatga	2160
tactggtaca	aaagctgggt	gctggtacaa	aagcagacac	atagatcaat	ggaacagaat	2220
agagggccca	gaaataaagc	tacacaccta	caaccatcta	atctttgaca	aagttgacaa	2280
aaatacgcaa	tggggaaaga	attccccatt	cagtaagtgg	tactgggata	actagctagc	2340
catatgcaga	ggattgaaac	tgaaccactt	ccttacacca	tatgcaaaaa	tcaactcaag	2400
atggattaaa	gacttaaatg	taaaaccccc	aactataaaa	actctggaag	ataacctagg	2460
caataccatt	ctggacatag	gaacggaaaa	agatttcatg	acaaagatcc	caaaaataat	2520
tgtaacgaaa	gcaaaaattg	acaaatggga	catgattaaa	cagaattacc	atttgactca	2580
gcaatcccat	tattggttat	ataccacaa	gaatctaaat	cattctgtca	taaagacata	2640
tatacacaaa	tgttcacggc	agcactatac	acaatcgcaa	agtcagggaa	tcaaactaaa	2700
tgtccatcag	tggtagaaag	gataaagaaa	atgtggtggc	agggagtggt	ggctcatgtc	2760
tgtaatccca	gcactttggg	aggctgaggc	gggtggttca	cctgagggtca	ggagtttgag	2820
accagcctgg	ccaacatggc	gaaactccgt	ctccgctaaa	aatacgaaaa	ttagccaggc	2880
gtggtggcga	gcacctgtca	tcccagctac	tgtggaggcc	taggcgtgag	aatcgcttga	2940
acctggaagg	tgggtgttgc	agtgcgcga	gatccctgcca	ctgcaactcca	gcctgggcaa	3000
ccaagcgaga	ctctgcctta	aaaaaaaaaa	aaagaaaatg	tggcacatat	acaccatgga	3060
atactatgca	gccataaaaa	agaatgggat	catgtcctgt	gcagcaacgt	ggatggagct	3120
ggaagccatt	atcctaaatg	aactcactca	gaaacagaaa	accaaatacc	acatgttctc	3180
acttataagt	agaagctaaa	cattgagtac	acatggatac	aaagaaggga	accgcagaca	3240
ctggggccta	cctgaggctg	gagcatggaa	ggagggtgag	gatcaaaaaa	ctacctatct	3300
ggtactatgc	tttttatctg	gatgatgaaa	taatctgtac	aacaaaacct	ggtgacatgc	3360
aatttaccta	tatatcaagc	ctacacatgt	gccctgaac	ctaaaaaaaa	agttaaaaga	3420
aaaacgtttg	gattattttc	cctctttcga	acaaagacat	tggtttgccc	aaggactaca	3480
aataaaccaa	cgggaaaaaa	gaaagggtcc	agttttgtct	gaaaaattctg	attaagcctc	3540
tggggccctac	agcctggaga	acctggagaa	tcctacaccc	acagaacccg	gctttgtccc	3600
caaagaataa	aaacacctct	ctaaaaaaaa	aaaaaaaa	3638		

<210> 262
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_017523

<400> 262
 ttggaaaagg aaaggtacta caaatccaaa agatttcact tttaacactg gcatttcctgc 60

<210> 263
 <211> 2461
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_018410

<400> 263
 atgctgggta cgctgcgcgc catggagggc gaggaacgtgg aagacgacca gctgctgcag 60
 aagctcaggg ccagtcgccg ccgcttccag aggcgcgatgc agcggtgat agagaagtac 120
 aaccagccct tcgaggacac cccggtggtg caaatggcca cgctgaoccta cgagacgcca 180
 cagggtattga gaatttgggg tggaaagacta ataaaggaaa gaaacaaagg agagatccag 240
 gactcctcca tgaagcccgc ggacaggaca gatggctccg tgcaagctgc agcctgggggt 300
 cctgagcttc cctcgcaccg cacagtcctg ggagccgatt caaaaagcgg tgagggtcgat 360
 gccacgtcag accaggaaga gtcagtgtgt tgggccttag cacctgcagt gcctcaaagc 420
 cctttgaaaa atgaattaa aaggaaatac ttgacccaag tggatatact gctacaagggt 480
 gcagagtatt ttgagtgtgc aggtaacaga gctggaaggg atgtacgtgt gactccgctg 540
 ccttcactgg cctcacctgc cgtgcctgcc cccggatact gcagtcgtat ctccggaaaag 600
 agtcctgggt acccagcgaa accagcttca tctcccagag aatgggatcc tttgcatcct 660
 tctccacag acatggcctt agtacctaga aatgacagcc tctccctaca agagaccagt 720
 agcagcagct tcttaagcag ccagcccttt gaagatgatg acatttgcaa tgtgaccatc 780
 agtgacctgt acgcagggat gctgcactcc atgagccggc tgttgagcac aaagccatca 840
 agcatcatct ccacaaaaac gttcatcatg caaaactgga actgcaggag gaggcacaga 900
 tataagagca ggatgaacaa aacatattgc aaaggagcca gacgttctca gaggagctcc 960
 aaggagaact tcataccctg ctctgagcct gtgaaaggga cagggggcatt aagagattgc 1020
 aagaacgtat tagatgtttc ttgccgtaag acaggtttta aattggaaaa agcttttctt 1080
 gaagtcaaca gaccccaa atccataagtt gatccaagtt ggaaggagcg caaagtgaca 1140
 cctcgaagt attcttctct gatttacttc gactccagtg caacatataa tcttgatgag 1200
 gaaaaatagat ttaggacatt aaaatggtta atttctcctg taaaaatagt ttccagacca 1260
 acaatacgac agggccatgg agagaaccgt cacagggaga ttgaaatccg atttgatcag 1320
 cttcatcggg aatattgcct gagtcccagg aaccagctc gccggatgtg cctcccggac 1380
 tctggggcca tgaacatgta cagaggggggt cctgcgagtc ctgggtggcct tcagggctta 1440
 gaaacccgca ggctgagttt accttcacagc aaagcaaaag caaaaagtgt aagtgaggct 1500
 tttgaaaacc taggcaaaaag atctctggaa gcaggtaggt gcctgcccac gagcgattca 1560
 tcttcatcac ttccaaaagac caaccccaca cacagcgcaa ctgcgccgca gcagacatct 1620
 gaccttcacg ttcagggaaa tagttctgga atatttagaa agtcagtgtc acccagcaaa 1680
 actctttcag tcccagataa agaagtgcca ggccaaggaa ggaatcgta cgatgaaatt 1740
 aaagaagaat ttgacaagct tcatcaaaaag tattgcctca aatctcctgg gcagatgaca 1800
 gtgcctttat gtattggagt gtctacagat aaagcaagta tggagttcg atatcaaaaa 1860
 gaaggcttct taggaaaaat aaatccagac cctcacttcc aggggtttcca gaagttgcca 1920
 tcatcacccc tggggtgcag aaaaagtcta ctgggtctca ctgcaattga ggctccttca 1980
 tctacatgtg ttgctcgtgc catcacgagg gatggcacga gggaccatca gttccctgca 2040
 aaaagaccca ggctatcaga accccagggc tcgggacgcc agggcaatc cctgggtgcc 2100
 tcagatgggg tggacaacac cgtcagaccg ggagaccagg gcagctcttc acagcccaac 2160
 tcagaagaga gaggagagaa cacgtcttac aggatggaag agaaaagtga tttcatgcta 2220
 gaaaaatttg aaactaaaag tgtgtagcta ggttatttcg gagtggtatt tatcttccca 2280
 cttgctctct gtttgtattt ttgttttgtt tttgattctt gagactgtga ggacttggtt 2340
 gacttctctg cccttaaagt aaatattagt gaaattgggt ccatcagaga taacctogag 2400
 ttcttgggtg agaaattatg tgaataaagt tgctcaatta gaaaaaaaaa aaaaaaaaaa 2460

a 2461

<210> 264
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_018410

<400> 264
 agtgatttca tgctagaaaa attggaaaact aaaagtgtgt agctagggtta tttcggagtg 60

<210> 265
 <211> 1405
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_018455

<400> 265
 cacctcgctc gcagcctccc cagcgcagca gcccggtgtt gggcctgcgg cagccgggtc 60
 ttcttggtcc ccacctcctg gggccgacgg gcggcaggaa ggggctcggc gggacgcgcc 120
 gtcagggacc tgaggaggaa caacggaacg cgttcggaaac ggcttggaact cccgagactc 180
 acccgactcg tggccacacc gggagaactg aagcggcagt agccggcggg gacgcccagc 240
 ccgaaggccg gctgctaggg agcagacagc tgaaccgctt gccagacgcc gaaacccagt 300
 gacgcccctc accgctccac cgtgctcccc gctccccgcc cccgccgccc gcggggccca 360
 agggcgcatgc gccgcctgtc ctggaggggc ccatttccgt ccgtcgtggg gggaggcaca 420
 gtgagtcacac tggggcacgg cagcgtctaa gccacaagcc gagcacataa gccaggtcct 480
 aacggagcct atgtgtaagt ccaactactgg tgcaagggtt cactcttcta agaagagcgg 540
 cgtggggggc tgggcgacct tcgcttcagt cgctcccccg tgcagtcccc tgtgcccaag 600
 acacagcctg atgcttgtgc tccggtgggc ggagcttgga ggcggcggga actgcaattg 660
 gtggcttttga aggcgcggcg agcgggaaca gctcttgagg agtgagactg caggagatgt 720
 gggccgtgcc aaagagatgg atgagactgt tgctgagttc atcaagagga ccatcttgaa 780
 aatccccatg aatgaactga caacaatcct gaaggcctgg gattttttgt ctgaaaatca 840
 actgcagact gtaaatttcc gacagagaaa ggaatctgta gttcagcact tgatccatct 900
 gtgtgaggaa aagcgtgcaa gtatcagtga tgctgccttg ttagacatca tttatatgca 960
 attttcatcag caccagaaag tttgggatgt ttttcagatg agtaaaggac caggtgaaga 1020
 tggtgacctt tttgatatga aacaatttaa aaattcgttc aagaaaattc ttcagagagc 1080
 attaaaaaat gtgacagtca gcttcagaga aactgaggag aatgcagtct ggattcgaat 1140
 tgcttgggga acacagtaca caaagccaaa ccagtacaaa cctacctacg tgggtgacta 1200
 ctcccagact ccgtacgcct tcacgtcctc ctccatgctg aggcgcaata caccgcttct 1260
 gggtcaggag ttagaagcta ctgggaaaaa ctacctccga caagaggaga tcattttaga 1320
 tattaccgaa atgaagaaag cttgcaatta gtgaacatga aaggaaaata aaaattcctc 1380
 acagtcaaaa aaaaaaaaaa aaaaa 1405

<210> 266
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_018455

<400> 266
 ccgacaagag gagatcattt tagatattac cgaaatgaag aaagcttgca attagtgaac 60

<210> 267
 <211> 927
 <212> DNA

<213> Homo sapiens

<300>

<308> NM_018465

<400> 267

```

ggcagcggggc gaaaggagcc ggggcctgga ggtttgcgta ccggtcgcct ggtccccgca 60
ccagcggccgc ccagtgtggt ttcccataag gaagctcttc ttctgtcttg gcttccacct 120
ttaacccttc cacctgggag cgtcctctaa cacattcaga ctacaagtcc agaccagga 180
gagcaaggcc cagaaagagg tcaaaatggg gtttatattt tcaaaatcta tgaatgaaag 240
catgaaaaat caaaaggagt tcatgcttat gaatgctcga cttcagctgg aaaggcagct 300
catcatgcag agtgaaatga gggaaagaca aatggccatg cggattgcgt ggtctcggga 360
attcctcaaa tattttggaa ctttttttgg ccttcagacc atctctttta cagctggagc 420
gattaaaaaa aagaagccag ccttcctggt cccgattggt ccattaagct ttatcctcac 480
ctaccagtat gacttgggct atggaaccct tttagaaaga atgaaagggt aagctgagga 540
catactggaa acagaaaaga gttaaattgca gctgccaaga ggaatgatca cttttgaaag 600
cattgaaaaa gccagaaagg aacagagtag attcttcata gacaaatgaa atcatgctta 660
ccaatcaaat ctcaaagcac agaattattg acttgaatca tgggtttttac agttttttta 720
atgctcaaga ttttgatatt atagatttta ttttaaaata ttaaaatgca agatagtttt 780
gagctatttt aaaataaaat ttataacatt caacacaaaa tcatggaggt gctctaaata 840
acttttagat ttctctcttc tgtgtgcatt accaatatct aagtgtaaaa ttaataaatt 900
gttttgaatt cctggaaaaa aaaaaa 927

```

<210> 268

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018465

<400> 268

```

ggaacagagt agattcttca tagacaaatg aaatcatgct taccaatcaa atctcaaagc 60

```

<210> 269

<211> 1047

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018487

<400> 269

```

cccacttctc cagccagcgc cccagccctc ccgcccggcg ctgcaggtc ccgaggagcg 60
cagactgtgt coctgacaat gggaaacagcc gacagtgtat agatggcccc ggaggcccca 120
cagcacaccc acatcgatgt gcacatccac caggagtctg ccctggccaa gctcctgtc 180
acctgctgct ctgcgctgcg gcccggggc acccaggcca ggggcagcag ccggctgctg 240
gtggcctcgt gggtagtgca gatcggtctg gggatcttga gtgcagtcct aggaggattt 300
ttctacatcc gcgactacac cctcctcgtc acctcgggag ctgccatctg gacaggggct 360
gtggctgtgc tggctggagc tgctgccttc atttacgaga aacgggggtg tacatactgg 420
gccctgctga ggactctgct aacgctggca gctttctcca cagccatcgc tgccctcaaa 480
ctttggaatg aagatttccg atatggctac tcttattaca acagtgcctg ccgcatctcc 540
agctcgagtg actggaacac tccagccccc actcagagtc cagaagaagt cagaaggcta 600
cacctatgta cctccttcat ggacatgctg aaggccttgt tcagaacctc tcaggccatg 660
ctcttgggtg tctggattct gctgcttctg gcattctctg cccctctgtg gctgtactgc 720
tggaagaatg tcccaaccaa agggaaaaga gaccagaagg aaatgttgga agtgagtggg 780
atctagccat gcctctcctg attattagtg cctgggtgctt ctgcaccggg cgtccctgca 840
tctgactgct ggaagaagaa ccagactgag gaaaagaggc tcttcaacag cccagttat 900
cctggcccca tgaccgtggc cacagccctg ctccagcagc acttgcccat tccttacacc 960

```

ccttccccat cctgctccgc ttcattgtccc ctcttgagta gtcattgtgat aataaaactct 1020
catgtttattg ttcccaggaa aaaaaaa 1047

<210> 270
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_018487

<400> 270
aaccaaaggg aaaagagacc agaaggaaat gttggaagtg agtggaatct agccatgcct 60

<210> 271
<211> 2280
<212> DNA
<213> Homo sapiens

<300>
<308> U17077

<400> 271
ccgcccgcga ccagctacgc cccgtccgac gtgccctcgg gggctcgcgt gttcctcacc 60
atccctttcg ccttcttccct gcccgagctg atatttgggt tcttggtctg gaccatggta 120
gccgccaccc acatagtata ccccttgctg caaggatggg tgatgtatgt ctgcgtcacc 180
tcgtttctca tctccttgat gttcctgttg tcttacttgt ttggatttta caaaagattt 240
gaatcctgga gagttctgga cagcctgtac cacgggacca ctggcatcct gtacatgagc 300
gctgccgtcc tacaagtaca tgccacgatt gtttctgaga aactgctgga cccaagaatt 360
tactacatta attcggcagc ctggttcttc gccttcacg ccacgctgct ctacattctc 420
catgccttca gcatctatta ccactgatgc acaggcgcca ggccaagggg gaaatgctct 480
ttgaaagctc caattattgg tccccaaaag cagcttccaa cgtttgccat ctggatgaca 540
aacggaagat ccactaaaac gtccacggga ttaacagaac gtccttgacg actgagcgat 600
gacaccacac tttgttttggg cattttaaatt cactctgctg aataggagga agcttttctt 660
tttcttgga aaacaactgt ctcttggaat tatctgacca tgaacttgct cttctagaca 720
actcacatca aagccctcac tccactaatg gagaatccta gcccactaa tgccaagtct 780
gtttggggat tttgcctcag ctatgggctt ccctagagta ggtctagggg aatactcagt 840
ctgatctttt ttttgtttgt tttattttgt tttttttgag acggagtctc gctcttctc 900
caaggctgga gtgcagtgc gcatctcca ctactgcag gctccgcctc ccgggttccc 960
gccattctcc tgccctcagcc tcccgagtag ccgggactac aggcgcccac caccatgcc 1020
ggctaattta gttgtatttt tagtagagat ggggtttcac cgtattagcc aggatggtct 1080
cgatctctcg acctcgtgat ccgcccgcct ggccctccca aagtgctggg attacaggcg 1140
tgagccaccg tgcccggcct gattctctta aaattgaaga ggtgctgcca aggccttcag 1200
atctaaccga gatgcataga ccttggttct ggtacttggt cagcctgtgc tggggagccg 1260
tggtcccagag ttccctgagg ggctgacagg gtcaagccac cctgcccacc accctcccac 1320
ttccctctcc ctttctctct cagcattagg attcaaggga aatctgcatg aagccaattt 1380
tgagggtaga cgtgtgggga aaataaatca ttatacagta agacctgggg cttgaggggt 1440
ggggaatggg gaggggaagg catagcctgc tccctcatga gtctgacatc tcggaaactg 1500
agcagctgcc ggacgcctgg gtcaggaatc caagaccca cctcttaagg actggttctc 1560
cagaaagcac cctcagggga aaagggtgaaa acattacatc cgtggattct cctgccacaa 1620
ccgcatggga agaaaaggct gccgcaacat ctacagcagg agtgaaggac ccattgtccc 1680
ggaaccgcgc tgcgccacct gcactcacc ccctcacatt ctcttaagca cccgggtggc 1740
ctccgaggct ggcggaatgg tggtgcccac ggggttgggc aagggtcac caggacctca 1800
acgggcaaaag ttgtgcacac taaaatatca aatcaagggt cttggtttta aagtaaatgt 1860
ttttctaaag aaagctgtgt tcttctgttg acccagacga atagggcaca gccctgtaac 1920
tgcacgtgcc ttctgtcatt ggggaatgaaa taaattatta cgagaaaggg acttgctcta 1980
actggtttga ggccttacag ttttgtatct acatttttcc cctcctgggg tttgccccga 2040
caggacaga actacaggag tcatgggaaa gaaaattctg gcttcactac tgctcactgc 2100
tcactttctg atcactctga tacttttttt tttttttttt ttttgcaacc tgataccttg 2160
aaaagcttct atgtgtctct ccttttgttg cctggcagct gtctaggatg atcactgatt 2220

actattttact aagtagccac atgcaaataa aagttgtttg gtaaaatgga aaaaaaaaaa 2280

<210> 272
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> U17077

<400> 272
 tcaccaggac ctcaacgggc aaagttgtgc acactaaaat atcaaatcaa ggtgcttggt 60

<210> 273
 <211> 2554
 <212> DNA
 <213> Homo sapiens

<300>
 <308> X87949

<400> 273
 aggtcgacgc cggccaagac agcacagaca gattgaccta ttgggggtgtt tcgcgagtgt 60
 gagagggaag cgccgcggcc tgtattttcta gacctgccct tcgcctgggt cgtggcgcct 120
 tgtgaccccc ggcccttgcc gcctgcaagt cggaaattgc gctgtgctcc tgtgctacgg 180
 cctgtggctg gactgcctgc tgctgcccac ctggctggca agatgaagct ctccctgggtg 240
 gccgcgatgc tgctgctgct cagcgcggcg cgggccgagg aggaggacaa gaaggaggac 300
 gtgggcacgg tggtcggcat cgacttgggg accacctact cctgcgtcgg cgtgttcaag 360
 aacggccgcg tggagatcat cgccaacgat cagggcaacc gcatcacgcc gtcctatgtc 420
 gccttctact ctgaagggga acgtctgatt ggcatgccc ccaagaacca gctcacctcc 480
 aacccccaga acacggtctt tgacgccaa cggctcatcg gccgcacgtg gaatgacccg 540
 tctgtgcagc aggacatcaa gttcttgccg ttcaagggtg ttgaaaagaa aactaaacca 600
 tacattcaag ttgatattgg aggtgggcaa acaaagacat ttgctcctga agaaatttct 660
 gccatggttc tcaactaaaat gaaagaaacc gctgaggctt atttgggaaa gaaggttacc 720
 catgcagttg ttactgtacc agcctatttt aatgatgcc aacgccaaag aaccaagac 780
 gctggaacta ttgctggcct aaatggttat aggatcatca acgagcctac ggcagctgct 840
 attgcttatg gcctggataa gagggagggg gagaagaaca tcctgggtgt tgacctgggt 900
 ggccggaacct tcgatgtgtc tcttctcacc attgacaatg gtgtcttcga agttgtggcc 960
 actaatggag atactcatct ggggtggagaa gactttgacc agcgtgtcat ggaacacttc 1020
 atcaaactgt acaaaaagaa gacgggcaaa gatgtcagga aggacaatag agctgtgcag 1080
 aaactccggc gcgaggtaga aaaggccaag gccctgtctt ctcagcatca agcaagaatt 1140
 gaaattgagt ccttctatga aggagaagac tttcttgaga ccctgactcg ggccaaattt 1200
 gaagagctca acatggatct gttccggtct actatgaagc ccgtccagaa agtgttgga 1260
 gattctgatt tgaagaagtc tgatattgat gaaattgttc ttgttggtgg ctcgactcga 1320
 attccaaaga ttcagcaact ggttaaagag ttcttcaatg gcaaggaacc atcccgtggc 1380
 ataaaccag atgaagctgt agcgtatggt gctgctgtcc aggtcgtgtg gctctctggt 1440
 gatcaagata cagggtgacct ggtactgctt catgtatgtc cccttacct tggattgaa 1500
 actgtaggag gtgtcatgac caaactgatt ccaagtaata cagtgggtgc taccaagaac 1560
 tctcagatct tttctacagc ttctgataat caaccaactg ttacaatcaa ggtctatgaa 1620
 ggtgaaagac ccctgacaaa agacaatcat cttctgggta catttgatct gactggaatt 1680
 cctcctgctc ctctgtgggt cccacagatt gaagtcacct ttgagataga tgtgaatggt 1740
 attcttcgag tgacagctga agacaagggt acaggggaaca aaaataagat cacaatcacc 1800
 aatgaccaga atcgctgtac acctgaagaa atcgaaagga tgggttaatga tgcgtgagaag 1860
 tttgctgagg aagacaaaaa gctcaaggag cgcattgata ctagaatga gttggaaagc 1920
 tatgcctatt ctctaaagaa tcagattgga gataaaagaaa agctgggagg taaactttcc 1980
 tctgaagata aggagaccat ggaaaaagct gtagaagaaa agattgaatg gctggaaagc 2040
 caccaagatg ctgacattga agacttcaaa gctaagaaga aggaactgga agaaattggt 2100
 caaccaatta tcagcaaact ctatggaagt gcaggccctc ccccaactgg tgaagaggat 2160
 acagcagaaa aagatgagtt gtagacactg atctgctagt gctgtaatat tgtaaatact 2220
 ggactcagga acttttgtta ggaaaaaatt gaaagaactt aagtctcgaa tgtaattgga 2280
 atcttcacct cagagtggag ttgaactgct atagcctaag cggctgttta ctgcttttca 2340

ttagcagttg ctcacatgtc tttgggtggg gggggagaag aagaattggc catcttaaaa 2400
 agcgggtaaa aaacctgggt taggggtgtg gttcaccttc aaaatgttct atttaacaac 2460
 tgggtcatgt gcatctgggt taggaagttt tttctacat aagtgcacc aataaatgtt 2520
 tgttatttac actggtcaaa aaaaaaaaaa aaaa 2554

<210> 274
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> X87949

<400> 274
 aactttcctc tgaagataag gagaccatgg aaaaagctgt agaagaaaag attgaatggc 60

<210> 275
 <211> 1359
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig1632

<400> 275
 ttttaagaca gttacctgtt gtgctgctgt tacaatatat aatgaaacca agtcagggga 60
 gtgaatttat caatcttttg atgtaaagta aaaacgtagt tcacacttca ggagagaact 120
 tcatagcaca atgtctttct ataagatatt tttaatgatt tagtatttta caacatttgt 180
 ttaccatatt ttgatatacc atttttttct atctgcccag ttttattaaa aaaactatat 240
 attattttct aaagaaacaa tcatattttt atacaaaatt atgttttcag gtaacgaaat 300
 agatgtaggg tacagtggaa cataagcagt gttaccctcg gctgggagtc agtattatac 360
 aacaaatggt gagctggaac atgccctgtc tgtgctgtcc ctctgtgct gggtcgcgga 420
 tgtgtaggca acattgcctt atcacgctag gttcacctga cactttaaaa ggaaaaaaag 480
 ttccatagag ttctgtggtc acaaaattgt tttgctttta tcaaatactt taatagaacc 540
 aaagttgcag atattggaat gtatggaagt atctcagtct ctgcataaga ggattaaagt 600
 atggaaggat catttaatga ctgttttact tataagtcac taagtaatcc accatttctt 660
 atggatgatg cttaagcctg gtgaggtttg tactctaagg agcccagatc ataatgcagt 720
 gcatttcctt agcccttaga gtttcttgca aacattttaa aaaagacata ttttaagaaag 780
 aaagataaag aaaaaacata ttttaattact gtaaacaggt actgctttat gtttattttc 840
 tctctacttc aacaaaaatc agatctttga ggttttgctg acattgttgg tggttttgca 900
 catgttcttt ctaattggat ttatgaatag ttctatgggt tttcaaagat gaatcatgct 960
 aagaacactt ctgctttttg atccactgtt tgcagcagaa ttatatatat gtataggaaa 1020
 aatccacttt gaataatoca tgttttgtat ttggaaattg tttttaaaaa taaaaaggaa 1080
 aggaaatata taaagctgtt atttattctg catttcttac atatctatcg cttgtcagta 1140
 taccctgttt ggtatatatt gcctctgcac atctacattt gtatatgcaa cagtgcagct 1200
 tataatctaca taaactgtaa ataactcttt ctgtgaaagg atcatcatat caagatgata 1260
 ccaaaagtat gtaaaaagaa acctgcatta ttttgtaatt atttcttata gatatttcat 1320
 ggtaagatta gcagtcaata aagttacttt tttgccttt 1359

<210> 276
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig1632

<400> 276
 gggttttcaa agatgaatca tgctaagaac acttctgctt ttgatccac tgtttgcagc 60

<210> 277

<211> 994
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig3464

<400> 277
 tgaatgtata tattaagact gtagctgaat tgcacatgaa atcagattgc caacttcttg 60
 actttcaatg ttagacattt atccttaagt tgtgagcgat atatgtagca tgctgtgaaa 120
 tgtctgttat agctctttta ttcacagta ttaatacaga attatcattt gcgtttcttg 180
 gtacttttta ttcaatgtaa tcagaagctg tgatgttttg cctttgtagt cctgtgcttt 240
 gttactgtaa tttttttttt tttttttacg aagcacgtga ctggactaat gtaaggcaga 300
 tgacgtgac tttaagactg ctatatatat cagtctctta ctctataagg ttttaaatta 360
 gaataagctt ttatcaaata gataattgat gcaatttagg attcacgcaa gtttcagtgt 420
 caaatggcgg tcttatagtt tcaattctga aaatagcaaa cttataaacc agccacttta 480
 aacttggtct ggcaaaccag accctgctgt agatatagtc taaggtagtt aaccatataa 540
 gccttttcaa ctcttaatgc cctccacatg aatcagcagt taagaagggt ctagaaccca 600
 tgaaagcttt tgtatgtatt actagggttt gtttttctta tgtttgctga ttttacagtt 660
 ctgactaaag ctgacctaaa tggatcagtt tatgtgtaat attctagtgc tttaatgact 720
 ctttttttct ttggagggag ggtaacatta tttggacaga tgcagaagga actgtttagtg 780
 agtcaagaca aacacatctg aaataaagga actgtgtatt aacatgttaa caattcataa 840
 ctgcactttt tatgacattt tgaaaatcta tttataggta cagaacaatg ggttttgtta 900
 aactgtatca catttatact tgcagaaatt tatttcattg ttattagtag gaattttatt 960
 ggttcaataa aattggcaaa actgaacacc aaaa 994

<210> 278
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig3464

<400> 278
 ctgctgtaga tatagtctaa ggtagttaac catataagcc ttttcaactc ttaatgccct 60

<210> 279
 <211> 423
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig14683

<400> 279
 tatgttatgg atatcttatt ttagagtaag aatataaggc atagccatat ttatgaagggt 60
 agtaatactc tactaatcaa tacttagaag tttttgttat gactaatctg aatgcttttt 120
 agtttttcct taatctagtt atgttggtta tttataagtc agttttcaga ttaggaaaga 180
 aggtatttga ggggtgtcca tttccactga atagtaagat gatgcttact tagatttcca 240
 cagctgtttg aaagctctgt atttggctat aacggaaaac tttgttaggg atgcttgatg 300
 ttttgtgttt tgtttctaaa ggaagacagt gttttgttcc ttcttttagaa aacttgaaga 360
 atagaataat gagtccagga ttaatttggg ataaagtctt ttacttcata aattctgatt 420

ctg 423

<210> 280
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> Contig14683

<400> 280

aggaagacag tgttttgttc cttctttaga aaacttgaag aatagaataa tgagtccagg 60

<210> 281

<211> 391

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28552

<400> 281

atgccattga	tgtgaagaag	gtgtctgtgg	aagactttct	tactgacctg	aataacttca	60
gaaccacatt	catgcaagca	ataaaggaga	atatcaaaaa	aagagaagca	gaggaaaaag	120
aaaaacgtgt	cagaatagct	aaagaattag	cagagcgaga	aagactcgaa	cgccaacaaa	180
agaaaaagcg	tttattagaa	atgaagactg	agggtgatga	gacaggagtg	atggataatc	240
tgtctggaggc	cttgcagtcc	ggggctgcct	tccgcgacag	aagaaaaagg	acaccgatgc	300
caaaagatgt	tccgcagagt	ctcagtccaa	tgtctcagag	gcctgttctg	aaagtttcta	360
accatggtaa	taaaccgtat	ttataaattg	c			391

<210> 282

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28552

<400> 282

aagactttct tactgacctg aataacttca gaaccacatt catgcaagca ataaaggaga 60

<210> 283

<211> 450

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28947

<400> 283

ctcatccaag	gagctggggc	agacttcatt	gattctagag	agacctgttt	cagtgcctac	60
tcatccctgc	cctctgggtgc	cagcctcctt	accatcacgg	cttcactgag	gtgtaggtgg	120
gtttttctta	aacaggagac	agtctctccc	ctcttacctc	aacttcttgg	ggtgggaatc	180
agtgatactg	gagatggcta	gttgctgtgt	tacgggtttg	agttacattt	ggctataaaa	240
caatcttggt	gggaaaaatg	tgggggagag	gacttcttcc	tacacgcgca	ttgagacaga	300
ttccaactgg	ttaatgatat	tgtttgtaag	aaagagattc	tgttggttga	ctgcctaaag	360
agaaagggtg	gatggccttc	agattatacc	agcttagcta	gcattactaa	ccaactgatg	420
gaagctctga	aaataaaaaga	tcttgaacct				450

<210> 284

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28947

<400> 284

agacagattc caactgggtta atgatattgt ttgtaagaaa gagattctgt tggttgactg 60

<210> 285
 <211> 439
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig30875

<400> 285
 agaaatcaat gacagttgac aggaagagag gacgcataca acaggcaaaa gaggaatgcc 60
 cagcagtctt ggtccttgcg gtgcaatact ggccttgagg ccaagtcagc aggggattcg 120
 tagtcactaa cttctaactg aggcagggaa gtaccatggt ctggaaaagg tccaaagaaa 180
 caggaataga ggcagtgtag caagaggcag atttttgggtg ccaaatagat ttgaatcctg 240
 gttctgcttc ttcctttgta gagtatgata ttggttcttt cctcccaaag ctattataaa 300
 gactaaatat gtacacaaat ctttgggatg tctgacatat aaatgcttaa caatagggtat 360
 ttgctgggtat tattacaaat gaatttgctt atttttgagc cacttctatg tctgtccatt 420
 aaacccaaaat gtgttctgc 439

<210> 286
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig30875

<400> 286
 ggttctttcc tcccaaagct attataaaga ctaaataatgt acacaaatct ttgggatgtc 60

<210> 287
 <211> 338
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31221

<400> 287
 gggaagttac actgcttcac accacaaggc cgtgggaaat cttggagggtt ctgtgccttt 60
 ctgtcacctc tactttttgc agctgtgatt gcactgtccc gcacatgtga ctacaagcat 120
 cactggcaag gaccctttta atggtgaaaa tgggcagatg aatagcaata agtggacctt 180
 tgttactctt ctgagttaga aaaattctaa ttttagtacac tctgaacaaa gcttattata 240
 cttacttaag atgtgttttg atttgggtgt cagaaagcaa cctgacaatg ataatactgt 300
 aactatgata aaattgagaa taaaaagatt ttatttag 338

<210> 288
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31221

<400> 288
 aaatgggcag atgaatagca ataagtggac ctttgttact cttctgagtt agaaaaattc 60

<210> 289
 <211> 417

<212> DNA

<213> Homo sapiens

<300>

<308> Contig31288

<400> 289

```

gaatcacttg agcccgggag gttgaggctg cagtgagctg tgtttatacc actgcactcc 60
agcctgctgg gtaacagagc aagactccat ctcaaaaaga aaagaaaaaa tgctttgcta 120
cataatgagg ccaggcaaaa aaaaaaaaaaag tcctgtggaa atcatataga caaacatttg 180
caaagctgct actgccattg taccagtgtt aaaatgtgtt ctaccttgca tcttttactg 240

atttttatga cagattttat attgtaacca tttgagaact ctgtaagtgc tatggcttcc 300
ttaaactacg atttatcata tgctcccagt gtttactttg agactgaatg gcaaccagag 360
aatgtaaaca accaagggtgc atctgggtat gttttaaaat aaagattaat aaaagtt 417

```

<210> 290

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig31288

<400> 290

```

ggcttcctta aactacgatt tatcatatgc tcccagtgtt tactttgaga ctgaatggca 60

```

<210> 291

<211> 394

<212> DNA

<213> Homo sapiens

<300>

<308> Contig31646

<400> 291

```

gctgctacac cccatgtaaa aagcggaaaa taaaatgaag attttccagc gcaagatgcg 60
gtactgggtg cttccacctt ttttggcaat tgtttatttc tgcaccattg tccaagggtca 120
agtggctcca cccacaaggt taagatataa tgtaatatct catgacagta tacagatttc 180
atggaaggct ccaagaggga aatttggttg ttacaaactt cttgtgactc caacttcagg 240
tggaaaaact aaccagctga atctgcagaa cactgcaact aaagcaatta ttcaaggcct 300
tatgccagac cagaattaca cagttcaaat tattgcatac aataaagata aagaaagcaa 360
gccagctcaa ggccaattca gaattaaaga ttta 394

```

<210> 292

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig31646

<400> 292

```

gccagaccag aattacacag ttcaaattat tgcatacaat aaagataaag aaagcaagcc 60

```

<210> 293

<211> 357

<212> DNA

<213> Homo sapiens

<300>

<308> Contig37562

<400> 293

```
caattatttc aagtgcacct tattaacaaa agtatcagtg gatccaacat aaaattttat 60
agtactaaat gtcaagccta actgtgaatt ttgttctgta tcttaagtaa atttatgata 120
atgttctcga gctatcaaca aaatatatgt acttttgtga gctatgaatt ttctaattaa 180
attttacatg ctataacatg atttttacat gaatgatact ttgtttataa ctatcaaatg 240
tcagtatttt actacaattt tattataaag tgtacattat cactaaatga acttcgattt 300
taaaaatcaa attagcttta gttgtatatt attttttaca aataaagata gacttgt 357
```

<210> 294

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig37562

<400> 294

```
atcaaatgtc agtatttttac tacaatttta ttataaagtg tacattatca ctaaatgaac 60
```

<210> 295

<211> 351

<212> DNA

<213> Homo sapiens

<300>

<308> Contig37895

<400> 295

```
aatagagaca cctctaatta attaaagcgg atgccctccc cactcctccc aggatttgac 60
tcggagcaca aactcttcac aaacccaaat gtcaggacac catcgccagt gtccactggc 120
cactgctgtt ggtgtgaggc agccaggagc ccctcagaac tagtaagtct gagaagaggc 180
tgcacggggc ctaggagagg gagaaatgag cccgtccaag gtgaattcct tgattctoca 240
ttgtgagtgc accaagaaca agcactccct ccgactgact ctgcctacc aggatctgga 300
acaccttcca ttaattttatt cgttcattca ataaatattt attgactgac t 351
```

<210> 296

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig37895

<400> 296

```
ctctcgcccta ccaggatctg gaacaccttc cattaatttta ttcgttcatt caataaatat 60
```

<210> 297

<211> 418

<212> DNA

<213> Homo sapiens

<300>

<308> Contig38288

<400> 297

```
gacaagtaaa tggggggccgt tgggacggcg ggtgcctgga gggcagctct ggggtcagcg 60
ggcagtgtct agagcacagg cccctctgtt ggggggatggg gagggagagca gtctgcctt 120
gggagcgtag gcccagggg gacttctaaa gccccccctg tcgtctgctc ttcaccacgc 180
accacagagg cacctgctgc acacacaagc atctcactcg gccacggag gggggccaggc 240
```

ttccttttggc tgaagctggt ttgggaaggg tctccacaca ggcactgac tcccaagctt 300
 tggatcatgat gtcttttacc atttgataat tttaaacatt gtttttaaac ccaaacatt 360
 tagtgggtccg ttgcctctga agatgtaaac aaacaataac actatttctg ggaacatt 418

<210> 298
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig38288

<400> 298
 ttttagtgggc cggtgcctct gaagatgtaa acaaacaaat acactatttc tgggaacatt 60

<210> 299
 <211> 413
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig38901

<400> 299
 tacatttttg tttaatgttg ggccctgaggt taactgtgac catgggtccag cttgagtggc 60
 ttctggagca gccacatttt caaggactgt ccaaaagcca gccagttcag ggctcaggcc 120
 tcacccattg cccactcctg gggagaccat cacctggctc atcgtttcca ccaagagtgc 180
 cccacaggag tgccccacag acccgctgga ccagcctgct gggggctcctg gccaggggtc 240
 tggctaacgg tgagggctga ctctgaactg totctcagtc tccagaaagt gttcaagcct 300
 gttgtgttcc caaatctgat tctccttatt gtcttgtaaa tcaaactcta agtgaact 360
 tcccattgt ccttcaaag attttttttt attaaatggt tttttaagat cct 413

<210> 300
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig38901

<400> 300
 tgttcccaaa tctgattcct cctattgtct tgtaaataca actctaagtg aaaacttccc 60

<210> 301
 <211> 434
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40434

<400> 301
 gaatgggtgaa agagagatgc cgtgttttga aagtaagatg atgaaatgaa tttttaattc 60
 aagaaacatt cagaaacata ggaattaaaa cttagagaaa tgatctaatt tccctgttca 120
 cacaaaacttt acactttaat ctgatgattg gatattttat tttagtgaat catcatcttg 180
 ttagctaaact ttaaaaaaatg gatgtagaat gattaaaggt tgggtatgatt tttttttaat 240
 gtatcagttt gaacctagaa tattgaatta aaatgctgtc tcagtatttt aaaagcaaaa 300
 aaggaatgga ggaaaattgc atcttagacc atttttatat gcagtgtaca atttgcctggg 360
 ctagaaatga gataaagatt atttatattt gtccatatct tgtacttttc tattaaaatc 420
 attttatgaa atcc 434

<210> 302
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40434

<400> 302
 aaggaatgga ggaaaattgc atcttagacc atttttatat gcagtgtaca atttgctggg 60

<210> 303
 <211> 391
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40552

<400> 303
 caccaagccc tgctccggca cctcgaatcc ctggcgacca tgagtcacca gctccaagcc 60
 ttactgtgcc ccagaccaa gagctccatc cccgcgctc tgcagcgttt gtctagcgcc 120
 cttgcagctc cagagccccc tggcccagcc cgtgactcct ctttggggcc tacagatgaa 180
 gctggctctg agtgtccctt ccctagaaaag gcctgacctt ccttaccac cagaacaggg 240
 gttttgatgc cctcactagt gttgaagcct gttccagaga gaggtgggac tgcaaggaga 300
 ggatggtcag ccctaccac ctgccctgtt tgagcttcct gtttgacaat gtttgctgtt 360
 gattttttgt tcaataaaga atttggtaaa a 391

<210> 304
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40552

<400> 304
 tttgagcttc ctgtttgaca atgtttgctg ttgatttttt gttcaataaa gaatttggtg 60

<210> 305
 <211> 495
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig41413

<400> 305
 aaatatctt aatagggcta ctttgaatta atctgccttt atgtttggga gaagaaagct 60
 gagacattgc atgaaagatg atgagagata aatgttgatc ttttggcccc atttggtaat 120
 tgtattcagt atttgaacgt cgtcctgttt gttgttagtt ttcttcatca tttattgtat 180
 agacaatttt taaatctctg taatatgata cattttccta tcttttaagt tattgttacc 240
 taaagttaat ccagattata tggtccttat atgtgtacaa cattaaaatg aaaggctttg 300
 tcttgcatgt tgaggtacag gcggaagttg gaatcagggt ttaggattct gtctctcatt 360
 agctgaataa tgtgaggatt aacttctgcc agctcagacc atttcctaag cagttgaaag 420
 ggaaacaagt atttcagtct caaaattgaa taatgcacaa gtcttaagtg attaaaataa 480
 aactgttctt atgtc 495

<210> 306
 <211> 60
 <212> DNA

<213> Homo sapiens

<300>

<308> Contig41413

<400> 306

cagctcagac catttcctaa tcagttgaaa gggaaacaag tatttcagtc tcaaaattga 60

<210> 307

<211> 409

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41538

<400> 307

aaaaaaaaaa	aaaaaaaaaa	aaagagttgt	tttctcatgt	tcattatagt	tcattacagt	60
tacatagtc	gaaggtctta	caactaatca	ctggtagcaa	taaagcttc	aggccacat	120
gatgctgatt	agttctcagt	tttcattcag	ttcacaatat	aaccaccatt	cctgccctcc	180
ctgccaaagg	tcataaatgg	tgactgccta	acaacaaaa	ttgcagtctc	atctcatttt	240
catccagact	tctggaactc	aaagattaac	ttttgactaa	ccctggaata	tctcttatct	300
cacttatagc	ttcaggcatg	tatttatatg	tattcttgat	agcaatacca	taatcaatgt	360
gtattcctga	tagtaatgct	acaataaatc	caaacatttc	aactctgtt	409	

<210> 308

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41538

<400> 308

ctcatgttca ttatagttca ttacagttac atagtcgaa ggtcttaca ctaatcactg 60

<210> 309

<211> 552

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41887

<400> 309

ctgaagacta	cgaccatgaa	atcacagggc	tgcggtgtgc	tgtaggtctt	ctcctgggtga	60
aaagtgtcca	ggtgaaactt	ggagactcct	gggacgtgaa	actgggagcc	ttaggtggga	120
atacccagga	agtcaccctg	cagccaggcg	aatacatcac	aaaagtcttt	gtcgcttcc	180
aagctttcct	ccggggtatg	gtcatgtaca	ccagcaagga	ccgctatttc	tattttggga	240
agcttgatgg	ccagatctcc	tctgcctacc	ccagccaaga	ggggcagggtg	ctggtgggca	300
tctatggcca	gtatcaactc	cttggcatca	agagcattgg	ctttgaatgg	aattatccac	360
tagaggagcc	gaccactgag	ccaccagtta	atctcacata	ctcagcaaac	tcaccctgtg	420
gtcgctaggg	tggggtatgg	ggccatccga	gctgaggcca	tctgtgtggt	ggtggctgat	480
ggtactggag	taactgagtc	gggacgctga	atctgaatcc	accaataaat	aaagcttctg	540
cagaatcagt	gc	552				

<210> 310

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41887

<400> 310

tactggagta actgagtcgg gacgctgaat ctgaatccac caataaataa agcttctgca 60

<210> 311

<211> 745

<212> DNA

<213> Homo sapiens

<300>

<308> Contig42342

<400> 311

```

gcagtaaaga caggacgcac ccatgtcaca agaggagcac aggcaggggt gttggtgttg 60
gggcagccct caggggtctcc agaccagcc ccaactcacac agcagcctag gaaggaaggg 120
cagagtccca ggtgtcagct ggtgggtctc ccaggagctg cccctccctg gaagtcacag 180
gacaggaatg acagatcagg gaactgcagg aagctgccac ctctgggggtc agaatatgcc 240
cagcctgctg gggctctcta tcgggggtctt cgagagccag acagcctgcc ttgtgctgca 300
tacctggctt tgctctgtgc agaaccagc acacgtgatt ttgtgtgaca tgccagcagc 360
ctggctccca ggacaggagg cctgccctgg gggaggggct gcaggaggag ggggggcagg 420
caccatgag tctgtccagc cttgtcacag atgcctgcc caagctgctg tcctgatttc 480
agctcacctc agagtaaactc agaataaact gcaccagac tttaacgaat gcatgttgac 540
gctttcagtt cacccttttc ttgtctaact ttcttctat tttcttctaa tgcgagagct 600
tattaattcc atatttatca ttttgaataa cttttctct ttttagtaac aaaatgtact 660
tcactcttag taaaatgtat ttactatttt agtaacaaaa atatacttgc ctaatcatgt 720
ttaaaatata gtgatgtgaa aaatt 745

```

<210> 312

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig42342

<400> 312

caccagact ttcacgaatg catgttgacg ctttcagttc acccctttct ttgctaactt 60

<210> 313

<211> 398

<212> DNA

<213> Homo sapiens

<300>

<308> Contig43645

<400> 313

```

agttcaaagg cagataaatc tgtaaattat tttatcctat ctaccatttc ttaagaagac 60
attactccaa aataattaaa ttttaaggctt tatcagggtc gcatatagaa tcttaaattc 120
taataaagtt tcatgttaat gtcataaggat ttttaaaaga gctataggta atttctgtat 180
aatatgtgta tattaaaatg taattgattt cagttgaaag tattttaaag ctgataaata 240
gcattagggg tctttgcaat gtggtatcta gctgtattat tgggttttatt tacttttaac 300
atthtgaaaa gcttatactg gcagcctaga aaaacaaaac attaatgtat ctttatgttc 360
ctggcacatg aataaacttt gctgtggttt actaatct 398

```

<210> 314

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig43645

<400> 314

gaaaagctta tactggcagc ctagaaaaac aaacaattaa tgtatcttta tgtccctggc 60

<210> 315

<211> 478

<212> DNA

<213> Homo sapiens

<300>

<308> Contig44289

<400> 315

ctaaaaacaa cactcatcag tcttgggaaa tttgaacttt gatcaactta actaaagaag 60
gaagggtagt aagaatTTTT caaatacaaa tatttgccaa ttcacagatg ataacattta 120
aggccttcaa aagtaagggg ttttccttgt ttctccagtc agcttttgtc aactctaata 180
gttttttcat aaacattttt tatttgtata attgcaacag tttaagaaat tatcacaact 240
atttagaaac atttaaaatg ttctttttga tataagctat atacttggaa aaatacattg 300
gtatctaaaa tttgaggtgt gttaagactg ctttttgttt taaaaaatgg tttaacattca 360
aatttttgaa gtgttttatg cttcatatgg ctaagttgta gtttggcaga gttaacagca 420
taagaataaa catgctgtaa ttttaaaaga tgctttgaat aaaaatttat ttttaattt 478

<210> 316

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig44289

<400> 316

catcagtcctt gggaaatttg aactttgatc aacttaacta aagaaggaag ggtagtaaga 60

<210> 317

<211> 556

<212> DNA

<213> Homo sapiens

<300>

<308> Contig44909

<400> 317

accatctggg atttctacag cctgggtacc catagccaca ccaaggcttc tgggagattc 60
tgcagggtca gctttccagg ctgttcccaa atagctccct gcctccccac tgccccataa 120
gccacagcag aagagccatt catctcataa acaaaaagga agaggaaaga atgaggaagg 180
accctgtgca aggttatattg caggcagggg tgggcttgta cctgacagca cccacccctg 240
tgtggccccc aggccctcat caccctcaga cccctcctaa gcagttccct cattgctctt 300
tggtactaggc tgacagcagg aagagcaggg cccatgaccg ggtggaagtt cagttttggt 360
gtctgcttca agaggggggt ttacactctg attccaggac aagcactctg aggcgggtgg 420
gggagagaaa ccctggctct tcacccaggt ttcacacaca tgtaaataaa acactatgtt 480
agtatctaac acactcctgg atacagaaca caagtcttgg cacatatgtg atggaaataa 540
agtgttttgc aatctt 556

<210> 318

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig44909

<400> 318

tcacccaggt ttcacacaca tgtaaatagaa acactatggt agtatctaac acactcctgg 60

<210> 319

<211> 710

<212> DNA

<213> Homo sapiens

<300>

<308> Contig45032

<400> 319

```

aaagataggc ttctaagtta aggcaaatca ttcattctgt cattaaacia atacaaacca 60
ggcacctgtc atatgccaaag tgatattcaa aatggcccat gtagaccttt gtgaagtatg 120
tggcctaaca gacattaaac aaatgtctgt gaaactgaca taataaagta aggtaagtta 180
tatgtgagac attctctttt tataataatt cctgtaaaagc agtacttact taggtaatatga 240
tatcatactg ttttgtttta tttttttcct aagagctaaa acgtcatcct ctcttcagtg 300
atgtggactg ggaaaatctg cagcatcaga ctatgccttt catccccccag ccagatgatg 360
aaacagatac ctctattttt gaagccagga atactgctca gcacctgacc gtatctggat 420
ttagtctgta gcacaaaaat tttcctttta gtctagcctc gtgttataga atgaacttgc 480
ataattatat actccttaat actagattga tctaaggggg aaagatcatt atttaaccta 540
gttcaatgtg cttttaatgt acgttacagc tttcacagag ttaaaaggct gaaaggaata 600
tagtcagtaa tttatcttaa cctcaaaact gtatataaat cttcaaagct tttttcatct 660
atattttttg tttattgcac tttatgaaaa ctgaagcatc aataaaatta 710

```

<210> 320

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig45032

<400> 320

ttaacctagt tcaatgtgct tttaatgtac gttacagctt tcacagagtt aaaaggctga 60

<210> 321

<211> 726

<212> DNA

<213> Homo sapiens

<300>

<308> Contig46218

<400> 321

```

atacatattg ctttagagag caggtaggtg gccatgtggt cagcagtggt tccttaagaa 60
aataccatct ttctaagcca ctggaatttt tactttacta tttttaacat taatggatgt 120
caggatcatca acctcaagtc tttacatatc catgtatatt ccatatatat tgtttatata 180
ggcccaagtt tctccttaat tgggatctat atactaccag cacaacatca aaaacatgta 240
attgaatata tcagagctat atatgtaagg aaatgactgg tgaccccatc atcatcattg 300
ttgaattcat gttaagtaga ccctctaggg gaccataagg caattgagca cataacgaaa 360
aatgatgcaa taagaatgta tgcactctct ttgccaaatg catgtgcttt tgtgtaacgt 420
ggatgtaaac agaattgcag tgctgccgaa attcttgatc ttggctaaga gagtattttt 480
ccccttgtaa ttatgactct gagataaaat tgccattttg aaattttcaa agtaacaact 540
ttttttattt tatgaataaa cttgggattg caatttctct gatctgacaa tcaataactt 600

```


taacaaagat ctaaataagt gtttcaagga aagttttcct aagcaaagt aatattacct 660
 catttgggca tcattactct gtttaattcta tatcaaagga aataaacttg ctacttgcac 720
 taaatg 726

<210> 322
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig46218

<400> 322
 accataaggc aattgagcac ataacgaaaa atgatgcaat aagaatgtat gcaactctctt 60

<210> 323
 <211> 580
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig47096

<400> 323
 ggtgggtctct catccttgtg tgctgctctg ctaagagatg tccaaggcgg agccgggggca 60
 agatccttcc agactcatct gtcagagccc caagcccttt agaccagag cccaaggacc 120
 atgccttttg gacattagga ctgcagcctt tgcttctgtg tattttggag ttttgggtgac 180
 ttttgtcacc tggacacact catttggttag ccatagtggg ttcccttggg cagcaacagt 240
 gcatgtacct ctggatgtca tctgaggtga gaccaccgag gccttttctc tctgtgtaca 300
 gaggggagtt aggagttgct ggactggatg cattacgagg actggggaca gggtagaggg 360
 acatccaggg atcagggcat gagtgggggc aaccccccg cctctgccct ggcatgggtct 420
 ccgcatgggc tgaggtgtag ctgattggct gccacatttc ggccatgctg gctggcgtgc 480
 ccatgttgca gatattttcc cgagttcccc agaatggatg gtattgaatc tcagccacat 540
 gcaaacactgt gtccagcatt ctttgcaata aatacttttt 580

<210> 324
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig47096

<400> 324
 atattttccc gagttcccca gaatggatgg tattgaatct cagccacatg caaacactgtg 60

<210> 325
 <211> 632
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig47563

<400> 325
 gccatctagt ctgtgggttt ctgttgaagc agtctgaatt gactaaaaca gtcacttggga 60
 gtagttataa accactttcc tgttgaaagc agaacatgct gattcaactg ttttgttcaa 120
 tagcaatgat agattttgtt taagtcccct acactttctt atttctaaat gatcaagagt 180
 acacttcctg gcagtgatta aggagtgtgt atctaacaga aaaaatata ataccctgtg 240
 aacccgaata tggaattcag attgtttctg ccctcagat catacttaaa aaacaagcat 300
 acaaacaaac ataagggaac aaacagcaac cataacaaaa acaaaccttt aaagggtgggt 360

```

ttttgctgtg ataatgaat acggtactct gaaggagaaa aaagtttctc aaatgagctt 420
aaactgcaag tgatttaaaa attagagaaat ataattctta aagctattga aagtttcaac 480
cagaaaacct caagtgaatt ttgtatgtaa atgaaatctt gaatgtaagt tctgtgattc 540
tttaagcaaa caattagctg aaaacttggt attgtttgtag tttatgtagt aagtgacttg 600
gcacccatca gaaaataaag ggcattaaat tg 632

```

<210> 326

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47563

<400> 326

```
agcaaacaat tagctgaaaa cttggtattg ttgtagttaa tgtagtaagt gacttggcac 60
```

<210> 327

<211> 540

<212> DNA

<213> Homo sapiens

<300>

<308> Contig48913

<400> 327

```

accagagggg gtcccttttc cacagtaatg ggatcggtcg gtgtgccttc agggaggaag 60
agggaggtgg tcaagcttga aaaactggct ttaggatggg tctgactttg ttctccctcc 120
ccaagtgttc tcaacctcca ttctgcagtg ttcagagttt tagggaaagg gtttgggtgc 180
cccagcatcc aggtgtttgt tggcttagcg catgtgaagt gaaaaccttc tgggggttgtt 240
tggaagcagc tttctgggttc ttgtgattgt atcctgaggt ccagaaccc tattctccca 300
cgaggatcct cagtgaacct ggtggccaca cgcctggcca gcctgctggc tcctgggtga 360
gctgaagaac cttgcctgtg gcacttttcg aggggtgagct ggaaccgaga gaacatggtc 420
cccgctgctg gactcatgcg ggtcatttcc tgccggcctg gtttcgcctg gtcgtgtctt 480
tatgagcacc atgtaagcct ccttgtattg agataattgg gcattaaaca ttaaactgca 540

```

<210> 328

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig48913

<400> 328

```
tatgagcacc atgtaagcct ccttgtattg agataattgg gcattaaaca ttaaactgca 60
```

<210> 329

<211> 534

<212> DNA

<213> Homo sapiens

<300>

<308> Contig49169

<400> 329

```

cctaattgta acatttttta aaatacatat ttgggactct tattatcaag gttctaccta 60
tgttaattta caattcatgt ttcaagacat ttgccaaatg tattaccgat gcctctgaaa 120
aggggggtcac tgggtctcat agactgatat gaagtcgaca tatttatagt gcttagagac 180
caaactaatg gaaggcagac tatttacagc ttagtatatg tgtacttaag tctatgtgaa 240
cagagaaatg cctcccgtag tgtttgaaag cgtaagctg ataatgtaat taacaactgc 300

```

```

tgagagatca aagattcaac ttgccatata cctcaaattc ggagaaacag ttaatttggg 360
caaatctaca gttctgtttt tgctactcta ttgtcattcc tgtttaatac tcaactgtact 420
tgtatttgag acaaataggt gatactgaat tttatactgt tttctacttt tccattaaaa 480
cattggcacc tcaatgataa agaaatttaa ggtataaaat taaatgtaaa aatt 534

```

```

<210> 330
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig49169

```

```

<400> 330
catacacctc aaattcggag aaacagttaa tttgggcaaa tctacagttc tgttttttgct 60

```

```

<210> 331
<211> 602
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig49388

```

```

<400> 331
tgtcagtggg ggggtctctg cagccaaactg agactatctt gctgtgcctt gagccttcct 60
agggttttag agaacagcat tcaaaattcc ccgtcctgtc agtggttgcc ttgcacctc 120
ctcccctaaa gcagcgcggg gggcaaataa gacccacccc ctccctgcag cttcacaggg 180
acgtttcctt ccctccccgc aaccacccca ggctcccttg ggaggctgca gttgtggtac 240
acgtccccgg tgctggggtg gccgtgactc gggggcgggg cgatcgggtc tcagcccctg 300
ccttccccag tctctgggtc acccgaattt tcccacccct gcttctcccc gaggaggttg 360
agctcttgag caagttggga cttgggcccgg ggccctggaag aatgattggc tgggaggccg 420
cgggaggggg gccaggaggc ccggaccagt tgggaggagt gagcaggccc cgggggaggg 480
ggatgagcgc agtttgctcg ctttctctcc ctgccggccc cctccgcccc cacacacact 540
cgggacgtct tcattgaaga ttcacttaca aaggaatgtt tcaactaaata aaagaaaacc 600
ag 602

```

```

<210> 332
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig49388

```

```

<400> 332
cgggacgtct tcattgaaga ttcacttaca aaggaatgtt tcaactaaata aaagaaaacc 60

```

```

<210> 333
<211> 562
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig50728

```

```

<400> 333
gcgaatttgg gccccttgat cctctgatgg gagctgaaag gatgagaggt gggcatctag 60
atltaggggg gctgttcagg ctttgcaggc cccttacctg aacacataga aaccctggag 120
ctgtgactgt gtccatgtgt gtgtgtttgt ctgtgtgtgt tgcgggggat gggcacctgc 180

```

```

atgaatgtgg tagagaaaat ggctctgctc agaggggaaga tacgcatagc aaggcagggga 240
ccagagggaat cacaggcgcc tggagagcag ccggggaccg cctccagggga cctgccggct 300
tccctcagtc ctccaggggc ccagcactct tcccttaggc cctgtgagcg tcccttgta 360
ggatacattc tctcattttg ctgaagctga tttgattggg tgtctgtttc tcgcagccaa 420
aagagctctg aatgaggaaa gtgcttctgt gctaactccc cgcgtctcct gaatttcagt 480
cattcatgta cccgcctcga aatttttgca atatctgtgt accaactgtc catttactta 540
ataaagaagt tttcttttaa tt 562

```

<210> 334
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig50728

<400> 334
 tttgattggg tgtctgtttc tcgcagccaa aagagctctg aatgaggaaa gtgcttctgt 60

<210> 335
 <211> 400
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AI497657

```

<400> 335
tttttttttt tgcacttatg gtattttattg ttggaagatt gagtacctta atgcacacca 60
atgctcagat gacttggggg cacatagggg actgctgtca ccatgcctca ctctgcagg 120
gaaggggctg cctactataa accccagcgg gccagtgct gtgtccagaa caggtcctta 180
tattactgca gccacaatg gaactactga gtaggagcca aaagaggagg gagcaggaag 240
aggtggcatt tggagagggg agaccgcacc cacaggctctg ccacagcgcg tcaacgggat 300
gggtacttt tacagtcaag ttgacttcgg tgtccgccca ccactacct ttgtaggacc 360
actgaaacaa gggacatcca ccacggccca cagccggggc 400

```

<210> 336
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AI497657

<400> 336
 gagcattggg gtgcattaag gtactcaatc ttccaacaat aaataccata agtgcaaaaa 60

<210> 337
 <211> 475
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig50950

```

<400> 337
ctggaagagg ctcccaaccc agagtgtccc tgtgggaggc aggcagaagg tgacaattga 60
cacgatttcc tgcacgcgtc ctctctacc ttggaagcag ttagaatcta ccaggcacag 120
atgaggccgc ccttgccctga cggagcttga tgagcagccc ttggctctccg gttccaggac 180
tgagagccca gctgcctctg cccacccttc cccaggcctc tgccagcctc tggctgcacg 240
gtcaggccct gccccatggc aggcctgcca gagcttggct ggggaccctc cccgcctctg 300

```

gctccctgat gggctggatg taacttgtgt cttctagccc ctttaaggagc ccaggtgttt 360
 taagggaatga attgggtcact gcatcttgta tcgattatgg ttctgagaaa agcaaataac 420
 acttttggct gcattaaaaa aagcatcata tataaaataa agaagatgaa ggtct 475

<210> 338
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig50950

<400> 338
 gtcactgcat cttgtatcga ttatggttct gagaaaagca aatatcactt ttggctgcat 60

<210> 339
 <211> 860
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig51660

<400> 339
 ggatggcaac cttcagctag actgcctggc tcaaggggtgg aagcaatacc aacagagagc 60
 atttggctgg ttccggtgtt cctcctgcca gcgaagttgg gcttccgcca agtgcagatt 120
 ctgtgccaca cgtactggga gcaactggga tcccagggtc aggtgcgtat gaggtctctt 180
 ggccaaagggt gccagaagtg ctccctgggtcc caatatgaga tgcctgagtt ctccctcgat 240
 agcaccatga ggattctgag caacctgggtg cagcatatac tgaagaaata ctatggaaat 300
 ggcattgagga agtctccaga aatgccagta atcctggaag tgcctctgga aggatcccat 360
 gacacagcca attgtgagga atgcactttg ggcattatgtg gacagggctt aaaaagctac 420
 atgacaaagc cgtccaaatc cctactcccc cacctaaaga ctgggaattc ctcacctgga 480
 attggtgctg tgtacctcgc aaaccaagcc aagaaccagt cagatgagga aaaagaggct 540
 aaggggagtg ggtatgagaa attagggccc agtcgagacc cagatccact gaacatctgt 600
 gtctttatatt tgcctgctgt atttattgta gtcaaattgt ttacatcaga atgatgaaaa 660
 taggcttgcc actttctctt attttaattc catggtagtc aatgaactgg ctgccacttt 720
 aatataactg aaaattcatt ttgagaccaa gcaggatcaa gttttagtaa taaacactgg 780
 tttcctagcc atcctctgaa aacagtatga aacatgacca agtacataat ggatttagta 840
 ataaatattg tcgaattgct 860

<210> 340
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig51660

<400> 340
 gctgcttgta tttattgtag tcaaatgctt tacatcagaa tgatgaaaat aggcttgcca 60

<210> 341
 <211> 608
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig52490

<400> 341
 atcgtggcta gcggacagac acgagcctct tgggaatacc ttgtccatca cgtcatggcc 60

```

atgggtgcct tcttctccgg catcttttgg agcagctttg tcgggtggggg tgtcttaaca 120
ctactggtgg aagtcagcaa catcttctctc accattcgca tgatgatgaa aatcagtaat 180
gcccaggatc atctctctta ccgggttaac aagtatgtga acctgggcat gtactttctc 240
ttccgcctgg cccctcaggc ctacctcacc ctttcttct tgcgttatgt gaaccagagg 300
acctggggca ccttctctgt gggatatcctg ctcatgctgg acgtgatgat cataatctac 360
ttttcccgcc tctctcgtct tgacttctgc cctgagcatg tccccaagaa gcaacacaaa 420
gacaagtctt tgactgagaa ctgagtgagg ggcacagagc ctggggacaac aaaaacggac 480
aaggccagaa acagcttcat atggacactg ggacttagcc ccaagcctgg gtgtcctctg 540
aggccagcct ctccaccttc tgagcctgcg cccacactat tgaaaacact aatgaaagta 600
ctcctctg 608

```

<210> 342

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig52490

<400> 342

```
ccaggatcat ctctcttacc ggggttaacaa gtatgtgaac ctgggtcatgt actttctctt 60
```

<210> 343

<211> 1282

<212> DNA

<213> Homo sapiens

<300>

<308> Contig53598

<400> 343

```

catgccagca cctttgaacc ggtctcttag aagaagacac acatcctggg tgtacagtgg 60
tgaaatgggg agtgggtgcc cattctgaaa aacgaggcat tcctgctcat tccctctgct 120
tagctggtgg gcaggggaga gagggaaatg ccaaaaactt ggagtgaagg atgatgctat 180
tttttatttt taaatatatc ttcaggttat tttcttactg ttgcttcaga tctaattgta 240
aaggcagatg tccctctctc tccaccccg acgctgacct cggcctcagt cacggctctt 300
tgcatgatca cagtctctgtg ttctggcctg tggcagggcc gggaagggcc gctggcttcc 360
gaacagacgt ggttgctctc cacgagggcg atggggagcc cgcgggccct aagctttgtc 420
gcagatgtca tcattggcag aattacttgt ctgaaaaaat aagtagcatt gctgaaacac 480
acaaccgaat tctctacgat ggccatttgc tcattgtctt tctctgtgtg gtagtgagtg 540
accttgccag tgtttgcctg ctcagagtgg cccctcagaa caacagggct ggccttgga 600
aaaccccaaa acaggactgt ggtgacaact ctggtcaggt gtgatttgac atgagggccg 660
gaggcgttg ctgacggcag gactggagag gctgctgccc cggcactggc agcgaggctc 720
gtgtgtcccc caggcagatc tgggcacttt cccaaccag gtttatgctg ctccagggaa 780
gcctcgggtg cagagtgggt ggcagatctg accatcccca cagaccagaa acaaggaatt 840
tctgggatta cccagtcccc cttcaaccca gttgatgtaa ccacctcatt ttttacaatt 900
acagaatcta ttctactcag gctatgggcc tegtctcac tcagttattg ogagtgttgc 960
tgtccgcatg ctccggggcc cagtggtctc ctgtgctcta gatcatggtg actccccgc 1020
cctgtggttg gaatcgatgc cacggattgc aggcacaaatt tcagatcggtg tttccaaaca 1080
cccttgctgt gccctttaat gggattgaaa gcacttttac cacatggaga aatataat 1140
taatttgatg tgcttttcta caaggtccac tatttctgag tttaatgtgt ttccaacact 1200
taaggagact ctaatgaaag ctgatgaatt ttcttttctg tccaaacaag taaaataaaa 1260
ataaaagtct atttagatgt tg 1282

```

<210> 344

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig53598

<400> 344
ccactatttc tgagtttaaat gtgttttccaa cacttaagga gactctaataa aaagctgatg 60

<210> 345
<211> 601
<212> DNA
<213> Homo sapiens

<300>
<308> Contig53641

<400> 345
tggaggctgt ggatgatgct ttcaagacaa tggatgtgga tatggccgag gaacatgcca 60
gggcccagat gaggggcccag atgaatatcg gggatgaagc gctgattgga cgggtggagct 120
gggatgacat acaagtcgag ctccctgacct gggatgagga cggagattttt gccgatgcct 180
gggccaggat cccctttgct ttctggggcca gataccatca gtacattctg aatagcaacc 240
gtgccaacag gaggggccacg tggagagctg gcgtcagcag tggcaccaat ggaggggcca 300
gcaccagcgt cctagatggc cccagcacca gctccaccat ccggaccaga aatgctgcca 360
gagctggcgc cagcttcttc tcctggatcc agcaccgttg acgaactgca gcgatcttac 420
tggccaagcc agagcgccctc ctctcagatt ccttctcgac acagcacctt aggcggcttc 480
ttcctgtcag tccggaggtgg catgcaagat gaagctctct ttgctcttcc tgctttcatt 540
ttgtgctttt ccttgtgttt tcatgttttg ggtatcagtg ttacattaaa gttgcaaaat 600
t 601

<210> 346
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig53641

<400> 346
ctttcatttt gtgcttttcc ttgtgttttc atgttttggg tatcagtggtt acattaaagt 60

<210> 347
<211> 751
<212> DNA
<213> Homo sapiens

<300>
<308> Contig54242

<400> 347
aattactcaa agaaggagcc atttcagtta actcaagtga atgaaagact tttggaatct 60
gcagtgggtc cttccctggt gaccatttgg taacttgtaa tctgacccaa aactcttgag 120
ctgcaacagg ccttgccaga gggctcagga tgggaaagga agaaggggat aggaaaagaa 180
gaggtaattt tacatttccc ctttaaagta aatttttagcc aactcatcat tctgaaatgt 240
ccctataaag aatgagtcga actagaccag aagccagcct actccttctt acatagcttc 300
tccaacaggg gtagcaatga cctgtccact tcaaacacag ataaggcctg ccatcctcat 360
tgggttaaagg cacacgtgag actttcagtg ggctctgctg agaaggaagg cagcccagga 420
gtcagggtatg caggcattgc attgtcagtg tctgctctca gaggtttacac attcaattgc 480
ttccaagggg gaatctcctg ctctgtgaat gctatcagac cccaaaggcc aaccttgggc 540
tgggtctatg tacgttcttc cgaagcactg atgatcaaaa ttgaagacac attcagaggt 600
ttgattggtt gagattaact ggtgtggtgg ttggtgtatg tatgttttat ttttatgtct 660
ttgtatgtag ttctacataa tgcaaatgt gctttctgat ggacaagacc tcataactgt 720
gattaatatc aataaaaagg ggatgttgtg g 751

<210> 348

<211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54242

<400> 348
 gttaaatttta gccaaactcat cattctgaaa tgtccctata aagaatgagt cgaactagac 60

<210> 349
 <211> 637
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54661

<400> 349
 ggcagtgatg tctatgttga gattaactta tgtattgagg aaaatttgaa gtttattttt 60
 tcgatgaata aggcgtgcaa atgatttagt atagattaat gacatctttt ttagaaatat 120
 taaagtgagt attcctcatt atgtcatcat ttctgataat tagagtgcta atttgaatgt 180
 tagataatgt ttccacatct atacctatct ctttctaggg cacttctgac cctggggcct 240
 ggggtagggc tttaggccac agtagtgtct gtgttaagtt cactaaatgt gtatttaatg 300
 agaaacattc ctatgtaaaa atgtgtgtat gtgaacgtat gcatacattt ttattgtgca 360
 cctgtacatt gtgaagaagt agtttggaat tttgtaaagc acaaaccata aaagagtgtg 420
 gagttattaa atgatgtagc acaaatgtaa tgttttagctt ataaaaggct ctttctattt 480
 tctatggcaa agactttgac acttgaaaaa taaaaccaat atttgattta tttttgtaag 540
 tatttaggat attattttta ataaatgatt gtccattatc aatataatag ttgtgaaatg 600
 atttaagttaa ataaacttta tgcttctgtg tctgttg 637

<210> 350
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54661

<400> 350
 ctgtacattg tgaagaagta gtttggaat ttgtaaagca caaaccataa aagagtgtgg 60

<210> 351
 <211> 924
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig55188

<400> 351
 ggcacaagta ccgcaagcgg gcactcatcc tgggtgtcact gctggccttt gccggcctct 60
 tcgccgccct cgtgctgtgg ctgtacatct accccattaa ctggccctgg atcgagcacc 120
 tcacctgctt ccccttcacc agccgcttct gcgagaagta tgagctggac cagggtgctgc 180
 actgaccgct gggccacacg gctgcccctc agccctgctg gaacagggtc tgctgctgag 240
 ggctgccctc tgcagagcgc tctctgtgtg ccagagagcc agagacccaa gacagggccc 300
 gggctctgga cctgggtgcc cccctgccag gcgaggctga ctccgcgtga gatgggttgg 360
 taaggcgggg tttttctggg gcgtgaggcc tgtgagatcc tgaccaagc tcaggcacac 420
 ccaaggcacc tgccctctctg agtcttgggt ctgagttcct aatatcccgc tccttgctga 480
 gaccatctcc tggggcaggg tccttttctt ccaggtcct cagcgtgctg tctgctggtg 540


```

ccttctcccc cactactact ggagcgtgcc cttgctgggg acgtggctgt gccctcagtt 600
gccccccaggg ctgggtgccc accatgcccc ttctctcttc tctcctacc tctgccctgt 660
gagcccatcc ataaggctct cagatgggac attgtgggaa aggctttggc catggctctgg 720
gggcagagaa caagggggga gacacaagta gacctcaggt agaacgacac tgggcggagc 780
caccocaggg cctgctccca gggagtgtct gaggcgcac aggcccgttt tttaccagtt 840
tatatcacgg tcttcatttt taaaagtaac gctaactttg tacggacgat gtctcatgga 900
ttaaataata ttctttatgg cagt 924

```

<210> 352

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55188

<400> 352

```

agtaacgcta actttgtacg gacgatgtct catggattaa ataatatctt ttatggcagt 60

```

<210> 353

<211> 699

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55353

<400> 353

```

tgattatgcc aagagctcta aacagaagtt tgagaaggta aaaattaagt tgtagtatct 60
gagttgtttt tattttcttc ctttggtggt tatgaaggta ttcataagaa ctttaatttc 120
aggggaaaaa atgcctgatt tgctattttt gacatttcct cgtctcttaa gaagtcagtt 180
aaatatgttt tcatagttta tattcctggt tcatagatta ctgtgaaaca tgtattttaa 240
cctatgaatt ataaaatagt atttagattc tagcgtgagt taaatagatt agtcatatat 300
cttttagatt tgtggatttg acatgtaaat tatgtgttgt gtataagtaa gttagttact 360
aaacatatgg catgggttatt gataaacttg ttgctatttt tttccaaatg ctatcagtgt 420
ttgtggactt ttaaaaatta gtttgaattt tggaatgttc tgtgataaaa tataatttca 480
actattttgt acattttaa atgccatggt gtatatgtct gtatttaaaa atgttgtaaa 540
tatctgcatt ttaagaatta tgaaagattt tcctcaaaaa tgacagaact ctccatactt 600
aattgtgaca cattataaga tatctgattt taagcttttg gattttgttc taaaaattaa 660
gtttaaacat gctgaaaatt ccataaaaaa aaaatttttg 699

```

<210> 354

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55353

<400> 354

```

taaaatagta tttagattct agcgtgagtt aaatagatta gtcatatatc ttttagattt 60

```

<210> 355

<211> 809

<212> DNA

<213> Homo sapiens

<300>

<308> Contig56503

<400> 355

```

gcatgtgaga tgagtgactg cccgtgaatg tgtccacagt tgagagggtg gagcaggatg 60
aggggaatcct gtcaccatca ataatacatt gtggagcgcc actctgccc aagcgccacc 120
tgggcgagaca gcatggagct ctccatggcc aggcctgacct tgtgcatggt cccgtgtctgg 180
tgccccctttg cccgcctcct gcaaacctca caggggtcccc acacaacagt gccctccaga 240
agcagccccct cggaggcgaga ggaaggaaaa tgggggatggc tgggggtctc tccatcctcc 300
ttttctcctt gccttcgcat ggctggcctt cccctccaaa acctccattc cccgtgtgct 360
agccccctttg ccatagcctg attttgggga ggaggaagggt gcgatttgag ggagaagggtg 420
agaaagctta tggctgggtc tggtttcttc ccttcccaga ggggtcttact gttccagggt 480
ggccccagggt caggcagggt ccacactatg cctgcgcctt ggtaaagggtg acccctgcca 540
tttaccagca gccctggcat gttcctgccc cacaggaata gaatggagggt agctccagaa 600
actttccatc ccaaaggcag tctccgtggt tgaagcagac tggatttttg ctctgccccct 660
gacccccctgt cctcttttga gggaggggag ctatgctagg actccaacct cagggactcg 720
ggtggcctgc gctagcttct tttgatactg aaaactttta aggtgggagg gtggcaagggt 780
atgtgcttaa taaatcaatt ccaagcctc 809

```

<210> 356

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig56503

<400> 356

```

gaaaactttt aaggtgggag ggtggcaagg gatgtgctta ataatcaat tccaagcctc 60

```

<210> 357

<211> 976

<212> DNA

<213> Homo sapiens

<300>

<308> Contig56678

<400> 357

```

gaaggatata ctttgttata acttattatt ttgttctctg taaatacaag atgtttatag 60
gaaatatgta ttctgaactc tatctgcaga atgagtcact acacccaaat agttctatta 120
tttagaatgt gttaatttta aagggaacct ataggtatatt attacatat gcgatccaca 180
tttgtgtgaa agcatgtgat catactaacc cagcctcctg gaatgtcgct gtacgatgat 240
tgatgtcttt ttctcagtc ctagttacaa ttgttttagta tgctaatacag tccagttccc 300
tgaggtttaa gatcaaatat aaattactct gctttctogac tcattcagggt agcattgtac 360
ctgaacctga ttgctacttt ttcatcttaa atattatatt tctcatctta atctgccttc 420
ccctcatcca cagacatttg gagaaggaaa tgggagggtg tctgttatcc ctttctcttt 480
gctttgtccc cgttgttaga ctggcagcgt cagttgctcg gtgggcttgg ttagagccgt 540
gggtgaggca ggtggctggc ggggacagggt agaggctgag aggggaagtgg tggcatttac 600
tgctctgaca cttccactgt cccgtgctggg gatgctgggg ccaaggcctg tggggcctgt 660
gaactgcaca gccaggagca aggaacccac taaatactcc gtcacctcca tgtcccctct 720
acagtgttaa attattacat aagcagggtga aaggtagaag gcgaattatg tgagtaaata 780
tggtctgttt tctcttcagc aaaaatgact atttttgtgt gtgactaatt tatttttatt 840
attgtaaaga tacaataaac cggttgaaat atctgctttg ttgacaagcg tgtgctttct 900
ctggccttat tcgcgttctg ttctcctgca aatagcgcct tctaaaaaga agagtcagac 960
aataaactgg ttgaaa 976

```

<210> 358

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig56678

<400> 358
tattacataa gcaggtgaaa ggtagaaggg gaattatgtg agtaaatatg gtctgttttc 60

<210> 359
<211> 1118
<212> DNA
<213> Homo sapiens

<300>
<308> Contig57584

<400> 359
agctgttgtg catccagagg tggaaattggg gcccggcatt ccctcctcgt cccgggctgg 60
cccttgcccc caccctgcaa ctcttggttg agatgggctc agccaagagc gtcccagtc 120
caccagcgcg gcctccgccc acaacaagca tctggctcga gtggcggacc cccgttcacc 180
tagtgctggc atcctgcgca ctcccatcca ggtggagagc tctccacagc caggcctacc 240
agcaggggag caactggagg gtcttaaaca tgcccaggac tcagatcccc gctctcctac 300
tcttggtatt gcacggacac ctatgaagac cagcagtggg gaccccccaa gccactggt 360

gaaacagctg agtgaagtat ttgaaactga agactctaaa tcaaattcttc cccagagacc 420
tgttctgccc ccagaggcac ctttatcttc tgaattggac ttgcctctgg gtaccagtt 480
atctgttgag gaacagatgc cacttgga ccagactgag ttcccctcca aacagggtgt 540
ttccaaggag gaagcaagac agcccacaga aaccctgtg gccagccaga gctccgacaa 600
gcccctcaagg gaccctgaga ctcccagatc ttcaggttct atgcgcaata gatggaaacc 660
aaacagcagc aaggtactag ggagatcccc cctcaccatc ctgcaggatg acaactcccc 720
tgccaccctg aactacgac agggtaagcg gccttcaccc ctaagtgaag atgttagtga 780
actaaaggaa ggagccattc ttggaactgg acgacttctg aaaactggag gacgagcatg 840
ggagcaaggc caggaccatg acaaggaaaa tcagcacttt cccttggtgg agagctaggc 900
cctgcatggc cccagcaatg cagtcaccca gggcctgggt atatctgtgt cctctcacc 960
cttctttccc agggatactg aggaatggct tgttttctta gactcctcct cagctaccaa 1020
actgggactc acagctttat tgggctttct ttgtgtcttg tgtgtttctt ttatattaaa 1080
ggaagtaatt ttaaagtgtt ctttaaaaag gtatatgt 1118

<210> 360
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig57584

<400> 360
aggaatggct tgttttctta gactcctcct cagctaccaa actgggactc acagctttat 60

<210> 361
<211> 859
<212> DNA
<213> Homo sapiens

<300>
<308> Contig63649

<400> 361
gtcgcagggt accagtgtgc ggagttcctg ttgccaagct gaaggtggcc ctgggcaggc 60
acagggtgtg tcatatcttc agccaacagg accatcctcc ggaggggccac ctctggggac 120
ttcctacggg aagagagtga cagatttggg gcttctgtgt gtttctgccg cttcagtggg 180
gccgctgcgg gagacagcgg gtggatcctc cagcagcctg tctgtgagc ctgccttctc 240
aagtctactg ttaaaatcag gaccgggtcg tgtccgagcc tacaggccct gtctccgctc 300
cccaggcctg caggagttag gggctgcacc tgctcgctgg agaggagag gcagatttag 360
tggaagcctg gcatggactc ggactggcct ttggaagctc cctgccctga cgggttgct 420
gtcaccactg cgaagtgagg cttggaggac ctgcacctga gaaaggctgt gtgtggtctt 480

```

gggtccacac ctgccagagc taacttactg ccagacggcg acttactgtg ggccaccctc 540
agtgaaccgg ggtgtcctca gctggcccta cagagcactt ctgtgctggg gatgagtagg 600
aactctgggc gaggagggtc ccagcgccgc ccctcgatac agccctgctc tgccctctgc 660
ccgtacttat accaggtggg atccctgccc tgcattgcct ggggattggc tgggcttggg 720
cacgccctgc tgtggaactg gatgttttca gggagcccag cctttcctca tgtcaacaca 780
gttcacaata tagttttcaa agtacagttt aaaactcaaa agtaaaacttt tcagcaactc 840
aaaaaaaaa aaaaaaaaaa 859

```

<210> 362

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig63649

<400> 362

```

cagcctttcc tcatgtcaac acagttcaca atatagtttt caaagtacag tttaaaactc 60

```

<210> 363

<211> 1170

<212> DNA

<213> Homo sapiens

<300>

<308> Contig63525

<400> 363

```

gccatggctc cctggggcgga gcgagcactc gcggtgaac ccgctgcgcg cgggtgtggct 60
cacgctgacc gccgccttcc tgcctgacct actgctgcag ctccctgcgc cgggcctgct 120
cccgggctgc gcgatcttcc aggacctgat ccgctatggg aaaaccaagt gtggggagcc 180
gtcgcgcccc gccgcctgcc gagcctttga tgtccccaag agatattttt cccactttta 240
tatcatctca gtgctgtgga atggcttcc tctttgggtgc ctactcaat ctctgttcc 300
gggagcacct tttccaagct ggcttcatgg tttgtcaga attctcgggg cggcacagtt 360
ccagggaggg gagctggcac tgtctgcatt cttagtgcta gtatttctgt ggctgcacag 420
cttacgaaga ctcttcgagt gccctcactt cagtgtcttc tccaatgtca tgattcacgt 480
cgtgcagtac tgttttgga tttgtctatta tgtccttgtt ggcctaactg tgcctgagcca 540
agtgccaatg gatggcagga atgctacata acagggaaaa atctattgat gcaagcacgg 600
tggttccata ttcttgggat gatgatgttc atctgggtcat ctgcccata gtataagtgc 660
catgttattc tcggcaatct caggaaaaat aaagcaggag tggctattca ctgtaaccac 720
aggatcccat ttggagactg gtttgaatat gtttcttccc ctaactactt agcagagctg 780
atgatctacg tttccatggc cgtcaccttt gggttccaca acttaacttg gtggctagtg 840
gtgacaaatg tcttctttaa tcaggccctg tctgcctttc tcagccacca attctacaaa 900
agcaaatttg tctcttacct gaagcatagg aaagctttcc taccattttt gttttaagtt 960
aacctcagtc atgaagaatg caaaccaggt gatggtttca atgcctaagg acagtgaagt 1020
ctggagccca aagtacagtt tcagcaaagc tgtttgaaac tctccattcc atttctatac 1080
cccacaagtt ttcactgaat gagcatgcag tgccactcaa gaaaatgaat ctccaaagta 1140
tcttcaaaga attaattact aatggcagat 1170

```

<210> 364

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig63525

<400> 364

```

ctcttaccgc aagcatagga aagctttcct accatttttg ttttaagtta acctcagtca 60

```

<210> 365

<211> 632
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig64688

<400> 365
 aagaatgcta agatgatttc agatatcgaa aagaaaaggc agcgtatgat tgaagtccag 60
 gatgaactgc ttcggttaga gccacagctg aaacaactac aaacaaaata tgatgaactt 120
 aaagagagaa agtcttccct taggaatgca gcataattct tatctaattt aaaacagctt 180
 tatcaagatt attcagatgt tcaagctcaa gaaccaaacy taaaggaaac gtatgattca 240
 tccagccttc cagctctgtt atttaaagca agaacacttc tgggagccga aagccatctg 300
 cgaaatatca accatcagtt agagaagctc cttgaccagg gatgagaaga gcagtctact 360
 aaaatgtgcc tataggaaga ctagtctcat gctgttacct tctgaaactg tacctttata 420
 aatcaattgt ttgcaaaga agttatggcc tacttagaat ctaaaatttg ttattcaaat 480
 taaatggctg tgaacaatgt taaatagcat cagtttgtcc aatagtttta aaggccataa 540
 tcatcttttc tggtaatat cttgagtaat tttaaaatgt tgacacctta atcggtccca 600
 ggtatgagcc ataataaact tgtaaaatta ag 632

<210> 366
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig64688

<400> 366
 ggctgtgaac aatgttaaatt agcatcagtt tgtccaatag ttttaaaggc cataatcatc 60